STRESS AND RELATED PHENOMENA IN BRAZILIAN (NATAL) PORTUGUESE

by

Silvia de Oliveira Segundo

Thesis Submitted to the Department of
Phonetics and Linguistics
in Partial Fulfilment of the Requirements
for the Degree of

Doctor of Philosophy

at the

School of Oriental and African Studies

University of London

1993

ABSTRACT

This thesis is devoted to the study of stress and related phenomena in the dialect of Brazilian Portuguese (BP) spoken in Natal. We assume a version of Metrical Theory where the principles of Government Phonology are incorporated.

An analysis of the assignment of stress to three grammatical categories is presented: verbs, nouns, and adjectives. The analysis offered is an attempt to prove that these three grammatical categories are assigned stress with a single set of parameters available in the Universal Grammar.

The analysis is divided in two parts. Part I consists of a discussion of various aspects of verbs. The insertion of clitics in the Future and the Conditional is discussed to some extent. Other aspects analysed include the relations of government expressed through vowel harmony, less obvious derivations where harmonized primarily stressed nuclei arise, and cases where the spreading of elements occurs.

Nouns and adjectives are analysed in Part II. Of the three patterns of stress present in BP nouns and adjectives, final and penultimate are discussed in depth. Antepenultimate stress is discussed only briefly. It is shown that in BP, verbs, nouns and adjectives can present either analytic or non-analytic types of derivation. It is also claimed that certain rhythmic processes operate in BP. An analysis of the diminutive suffix is presented. According to this, there are two lexical entries which are referred to as the "diminutive": -iñ- and -ziñ-. The Natal data show that the former is the non-analytic version of the diminutive and the latter is the analytic version of it.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank a number of people for their support during the time I was preparing this thesis.

First of all, I thank my supervisor, Jonathan Kaye, for his moral and intellectual support. Words cannot express my gratitude for his kindness and patience.

Jonathan Kaye and Monik Charette, please accept my deepest thanks for your help over the Christmas of 1989, when you kindly helped me get over some bad moments.

I thank Monik Charette and Edmund Gussmann for accepting to be my examiners.

I thank my parents with enormous gratitude for their moral and financial support during my studies, and for their prompt understanding in every situation and patience throughout my PhD life. I dedicate this thesis to them.

Fernando, thank-you especially for your moral support. I owe part of my perseverance in finishing this work to you.

Many friends who helped in the various stages of my studies deserve to be mentioned here: Augusta Cavaco, Thaïs da Silva, Yong Heo, Csaba Pogany, Sanjik Rhee, Yilmaz Vural, Shohei Yoshida, and Yuko Yoshida, who supported me so many times in the Department; Lina Neto, Patrícia Rondó, Shohei Yoshida and Yuko Yoshida (again), when I lived in the International Hall; Karla Couto, Fátima, Fernanda, Theo, and Zezé, my earlier friends from Finsbury Park who helped me so many times; and Caio, Flávia, Laércio (Lalá), and Zezé, thank-you for finding a space for me in your place and for your help in many ways. Lalá, thank you for being so helpful and kind to me.

Other people helped me in other ways while I was writing this thesis. Barbara Forbes and Susan Gillian, thank-you for correcting my English in the previous version of this thesis, and Angela Halwood, thank-you for making a few of my lines sound better in English. Phillip Backley, thank-you very much for your help in the final version of this thesis.

This work was supported by the Brazilian CNPq (Grant no. 20.0511/87) and by Helisom.

1.1 *Introduction.*

This thesis will attempt to account for the assignment of stress to verbs, nouns and adjectives and is organized as follows. In Part I we shall present an analysis of several aspects of verbs in Brazilian Portuguese (BP). Chapter 2 considers the assignment of stress to the verb structures, and we will propose that verb paradigms are stressed according to a set of parameters available in Universal Grammar (UG). A discussion of some aspects of verb morphology is also presented.

Chapter 3 deals specifically with the insertion of clitics in two verb tenses, the Future and the Conditional. The discussion focuses on the stress changes that occur when clitics are inserted.

Chapter 4 offers an in-depth analysis of the governing relations which become manifest through vowel harmony, and the interaction of these relations with metrical structure. The case of mid vowels is considered in detail.

In Chapter 5 we present a summary of some rather less straightforward verb derivations. Specifically, our attention will focus on (a) derivations resulting in harmonized nuclei which bear primary stress: (b) cases of spreading of elements in the derivation of the second person plural in most tenses: and (c) the derivation of the first person plural of the future tense.

In Part II we shall present an analysis of several aspects of nouns and adjectives. A discussion of some morphological aspects of nouns and adjectives in chapter 6 serves as the introduction to an analysis of stress assignment in these categories. It is proposed that, with respect to the assignment of stress, nouns and adjectives are subject to the same parameter settings that are required for verbs. It is also proposed that nouns and adjectives are lexically marked to coincide with the first strong node of a word domain when metrical structure is created via leftward

scanning.

Chapter 7 deals with structures involving analytic domains. It will be shown that, in Brazilian Portuguese, nouns and adjectives can present either analytic or non-analytic types of derivations. Unlike other theoretical frameworks, the analytic/non-analytic distinction is a property of affixes, not of processes. We discuss a few cases where stress clashes arise and propose a rhythm rule whereby the alternating pattern of strong and weak nodes is re-established.

In Chapter 8 we offer a unified account of the diminutive, analysing it as two distinct diminutive suffixes: -iñ- and -ziñ-. The former is the non-analytic version of the diminutive and the latter is its analytic counterpart.

Chapter 9 presents a brief discussion on some matters relating to the assignment of antepenultimate stress in nouns and adjectives. Our arguments and proposals are then drawn together in chapter 10.

1.2 *Theoretical Assumptions and the Notation*

In this thesis we assume a variation of Metrical Theory where all principles of Government-based Phonology (henceforth GB Phonology) are incorporated. In order to acquaint the reader with this formalism, we will sketch the most important aspects of these theories that contribute to the version of Metrical Theory assumed below.

With respect to Metrical Theory, most of the assumptions proposed by Halle & Vergnaud (1987) are adopted here. However, there are two exceptions to this. Firstly, the notation employed here deviates significantly from their work. In this thesis we use the arboreal type of metrical structure. The second deviation from the work of Halle & Vergnaud is the interpretation of cyclicity

with respect to how phenomena occur in the various types of cyclic structures. In this framework, analytic[1] domains refer to the fact that phenomena occur independently in each domain and that, when the derivation reaches the larger domain, the structure obtained from the application of processes in previous stages is maintained. There are four logical possibilities with respect to bracketings.

(1) a. Non-Analytic Morphology.

[A B] Domain: AB

b. Analytic Morphology.

i. [[A] [B]] Domains: A, B, AB

ii. 1. * [A [B]]

2. [[A] B] Domains: A, AB

((1)a) represents the case where the strings in question are treated by the phonology as unanalysable words. In phonological terms, this type of string behaves exactly like words with no internal morphological structure such as 'emolument' or 'hippopotamus'. The structures in ((1)b) represent the cases where the strings A and B both constitute independent phonological domains. This means that processes will apply to each individual domain and will modify each individual structure appropriately and exhaustively; after deletion of boundaries, no resulting structure can be undone. Of course, after deletion of boundaries the newly derived structure can trigger other processes that are sensitive to that new context, but the resulting strings from the previous domains will remain the same. Stress assignment is a typical example of a process that cannot be subsequently altered.

The third type of analytic derivation includes the cases where one string, A in the case of ((1)b.ii.1) and B in the case of ((1)b.ii.2), does not constitute a domain for the application of processes. Although ((1)b.ii.1) is logically possible, "this type of structure

[1]Which is the term used here in order to distinguish it from the term used by Halle & Vergnaud or in Lexical Phonology.

does not seem to be attested" (Kaye, to appear). According to the theory, a string which is not surrounded by an independent set of brackets will receive no stress.

These are the two major deviations from the work of Halle & Vergnaud. To our knowledge, this thesis constitutes the first approach to the study of metrical structure within this combined framework.

With respect to "syllabic" constituents or the organisation of segments into constituents, this operation follows from the principles and parameters available in the Universal Grammar as proposed in works such as Kaye (1989, 1990a,b, 1992, to appear) and Kaye, Lowenstamm & Vergnaud (1985, 1990) (henceforth KLV (1985) and KLV (1990), respectively). What follows constitutes a summary of the principles and parameters that determine the organization of segments into constituents. We will ignore, for the time being, the question of which segments can attach to each constituent (at the segmental level). The reader is referred to the works cited above for a detailed discussion.

Basically, in GB Phonology constituents are defined, as expected, with reference to relations of government. Since government per se implies at least the existence of a head (the unit that has the power to exist by itself), it follows that each constituent has to start off with a head (=governor). After that, constituents will gain their structure by application of two properties, Strict Locality and Strict Directionality, which are set out below:

(2) *Strict Locality*

No position must intervene between governor and governee at the P_0 projection.

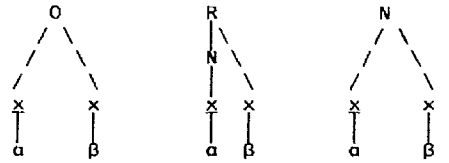
(3) *Strict Directionality.*

Directionality of government at the skeletal level is universal - the head is on the left.

P_0 is the projection on which all string positions are present (Kaye, to appear). It is stipulated that there are three syllabic

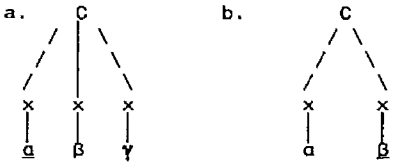
constituents: (O)nset, (R)hyme and (N)ucleus. It is also stipulated that the (N)ucleus is embedded in the (R)hyme. Since the construction of these phonological constituents depends on these two principles, it follows that they are maximally binary, as illustrated below[2].

(4)



Constituents such as the ones in (5) below are excluded since they violate the principles of Strict Locality and Strict Directionality, respectively. The constituent in ((5)a) is excluded because the head is not adjacent to the skeletal point dominating γ. The constituent in ((5)b) is excluded because the head is on the right, violating the Principle of Strict Directionality.

(5)



This exhausts all possibilities for constituent structure in all languages of the world.

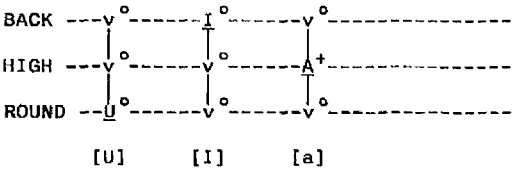
Before we go into the stress related aspects of the theoretical framework adopted in this thesis, we will present a short discussion on the representation of nuclear segments in GB Phonology which are especially relevant to the discussion on vowel harmony (chapter 4).

According to GB Phonology, the primary units of phonology at the segmental level are *elements*. These elements are univalent and independently pronounceable. Segments may consist of either isolated elements or a combination of them. There are five

[2]Heads are underlined.

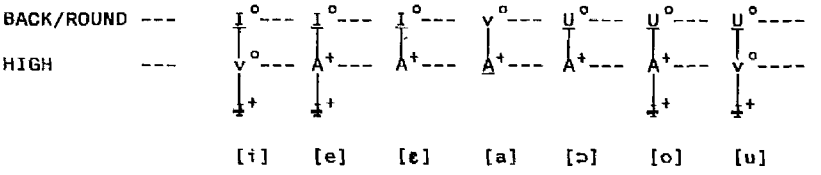
primitives for the combination of nuclear segments (=vowels), I°, U°, $A^{\dagger}$, v° and $\ddagger^{\dagger}$, the first three having "palatality", "roundness" and "pharyngeality" as their salient properties, respectively. These are the basic elements that can be involved in the composition of vowels, although they may be involved in the composition of consonants as well. The "salient" property of these elements is the contribution an element makes to an expression when it is an operator. KLV (1985) stipulate that the salient properties of I°, U° and $A^{\dagger}$ are [-back], [+round] and [-high], respectively, and they are individually pronounced [I], [U] and [a], respectively. KLV (1985) claim that in the absence of an element in a skeletal position there will always be a cold vowel, or v°, which has no salient property and is derived through the negation of the other three elements' salient property. The cold vowel is then [+back], [-round] and [+high]. It only manifests itself when it is the head of an expression, otherwise it has the same value as a zero on the left in numerical expressions. The remaining segmental prime which is of relevance to the discussion below is the $\ddagger^{\dagger}$ element. Its presence indicates tenseness of the compound (=segment), that is, the fact that the root of the tongue is pushed forwards, cf. the pair [ɛ]/[e] where the former is lax and the latter is tense. It is usually associated to the ATR (Advanced Tongue Root) feature of distinctive feature theory. Elements combine by means of an operation called *fusion*. The segment resulting from the fusion operation has all the properties of the head except for the salient property of the operator. According to KLV (1985), elements are located on *tiers*. Each tier is named after the salient property of the element. This being the case, the I° element is found on the BACK tier, the U° element is found on the ROUND tier, and the $A^{\dagger}$ element is found on the HIGH tier. These tiers are expressed through the following type of diagram, where some of the possible combinations of I°, U° and $A^{\dagger}$ are illustrated:

(6)



Lines may be conflated, so elements will be on the same line. This is subject to parametric variation. Since only elements found on distinct lines may combine, it follows that certain combinations are impossible if elements are found on the same line. In the particular case of BP, we claim that the BACK and ROUND lines are conflated. This entails that the combination of elements I° and U° is impossible, so rounded front vowels will never exist in BP. The vocalic system of BP can be obtained by the following combinations:

(7)



These combinations can also be expressed using the following type of expression:

(8)

$(\underline{\text{H}}^{\dagger}.\underline{\text{I}}^{\circ})^{\dagger}$	[i]
$(\underline{\text{H}}^{\dagger}.(\text{A}^{\dagger}.\underline{\text{I}}^{\circ})^{\circ})^{\dagger}$	[e]
$(\text{A}^{\dagger}.\underline{\text{I}}^{\circ})^{\circ}$	[ɛ]
$(\underline{\text{A}}^{\dagger})^{\dagger}$	[a]
$(\text{A}^{\dagger}.\underline{\text{U}}^{\circ})^{\circ}$	[ɔ]
$(\underline{\text{H}}^{\dagger}.(\text{A}^{\dagger}.\underline{\text{U}}^{\circ})^{\circ})^{\dagger}$	[o]
$(\underline{\text{H}}^{\dagger}.\underline{\text{U}}^{\circ})^{\dagger}$	[u]

The position which an element occupies in an expression determines the resulting segment. This means that whether an $A^{\dagger}$ combines as the head or the operator in an expression, for example, will produce distinct segments. As an example, the element $A^{\dagger}$ by itself (= as a head) gives [a], while its combination as an operator ($A^{\dagger}.\underline{\text{v}}^{\circ}$) gives [ə]. The expression yielding [a] can be understood as one where the resulting segment will have all the characteristics of the head except for the salient property of the operator. As mentioned earlier, the cold vowel only manifests itself when it is the head of an expression. Since the cold vowel

is [+back, +high, -round] and the salient property of $A^{\dagger}$ is [-high], the resulting segment should be [+back, -high, -round], which is a schwa. The expression where the roles are reversed yields [a] because the cold vowel has no salient property to contribute, as an operator.

Another aspect of GB Phonology which is relevant to the discussion that follows has to do with the property called *charm*. According to KLV (1985) and the update KLV (1990), elements can combine according to their charm[3]. An element can be positively or negatively charmed or it can be neutral. It is postulated that elements with like charm cannot combine. Just like elements, expressions will also be charmed or charmless. Whenever $A^{\dagger}$ is the head of an expression, it contributes its charm value to that expression. For example, the expression $(A^{\dagger}.\underline{v}^{\circ})^{\circ}$ is charmless because the head, the cold vowel, is charmless. The expression $(I^{\circ}.\underline{A}^{\dagger})^{\dagger}$, on the other hand, is charmed because the head, $A^{\dagger}$, is positively charmed. An expression (=a segment) can also combine with the elements $\text{ɪ}^{\dagger}$ or $N^{\dagger}$ to produce tense or nasalized segments, respectively. This combination is, of course, subject to the charm limitations postulated by the theory, according to which two identical charms cannot combine. This entails that $\underline{A}^{\dagger}$, for example, cannot combine with $\text{ɪ}^{\dagger}$ or $N^{\dagger}$ because both $A^{\dagger}$ and $I^{\dagger}$ and $N^{\dagger}$ are positively charmed. This also implies that the combination of the nasal element, $N^{\dagger}$, with a tense segment (a positively charmed expression) is excluded theoretically. The addition of the element $\text{ɪ}^{\dagger}$ is, therefore, only possible if the expression is neutrally charmed. What this means is that combinations like the ones in (9) below are theoretically excluded.

(9) * $(N^{\dagger}.(\text{ɪ}^{\dagger}.(A^{\dagger}.\underline{I}^{\circ})^{\circ})^{\dagger})^{\dagger}$ = [ẽ]
* $((N^{\dagger}.(I^{\dagger}.(A^{\dagger}.\underline{U}^{\circ})^{\circ})^{\dagger})^{\dagger})^{\dagger}$ = [õ]

It is also postulated that the elements $\text{ɪ}^{\dagger}$ and $N^{\dagger}$ always combine as operators. Nevertheless, it is postulated that their charm is

---

[3]The charm property is represented by a superscript on the righthand side of an element (Cf. previous page).

always transmitted to the resulting expression. The theory predicts, then, that the expressions in (10) below, for example, are perfectly admissible. This is indeed the case, as attested in many languages[4].

(10) $(N^{\dagger}.(A^{\dagger}.\underline{v}^{\circ})^{\circ})^{\dagger}$ = nasal schwa
$(N^{\dagger}.\underline{I}^{\circ})^{\dagger}$ = nasal lax high front vowel
$(N^{\dagger}.\underline{U}^{\circ})^{\dagger}$ = nasal lax round high back vowel
$(I^{\dagger}.(A^{\dagger}.\underline{I}^{\circ})^{\circ})^{\dagger}$ = tense front mid vowel
$(I^{\dagger}.(A^{\dagger}.\underline{U}^{\circ})^{\circ})^{\dagger}$ = tense round high (back) vowel

We will see in the vowel harmony analysis (chapter 4) what consequences these theoretical aspects entail in the analysis of verb paradigms in Natal.

One very important principle of GB Phonology is directly linked to the assignment of stress: the Licensing Principle. This principle is given below:

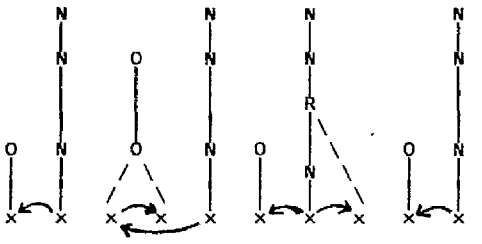
(11) Licensing Principle (Kaye (1990b)

All phonological positions save one must be licensed within a domain. The unlicensed position is the head of the domain.

According to this principle, in a domain, e.g. the word, every skeletal position must be licensed, except for the head of the domain. The complement of an onset head is licensed by the onset head. The head of the onset is licensed by the nucleus which follows and which governs it. Nuclear complements are licensed by nuclear heads and rhymal complements are licensed by nuclear heads. Governing heads are projected at the licenser projection as proposed by Charette (1990), and after that nuclear heads are projected at the Nuclear Projection. The licensing at lower projections up to the nuclear projection as just described can be illustrated with the following diagram:

[4]For a thorough discussion on precisely how this matrix calculus is done, the reader should refer to KLV (1985, 1990).

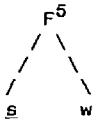
(12)



According to GB Phonology, it is at the Nuclear Projection that the metrical analysis is carried out. This means that nuclei with phonetic content only will be present at this level, since empty nuclei have already been licensed at the skeletal level.

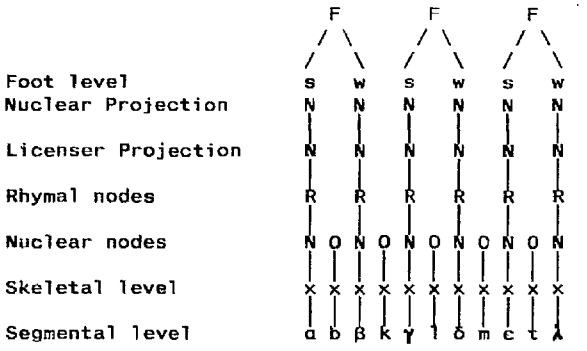
Foot construction is subject to two parameters. They can be left-headed or right-headed and they can be bounded or unbounded, according to the language.

(13)



(14) illustrates all levels of projection involved in the construction of metrical structure up to the foot level in BP.

(14)

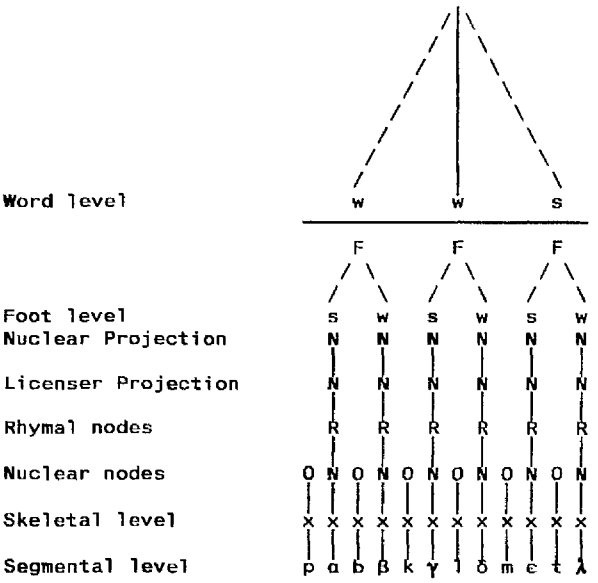


After that, feet are projected to an upper projection where the word tree is constructed (uppermost level of metrical structure in (15)).

---

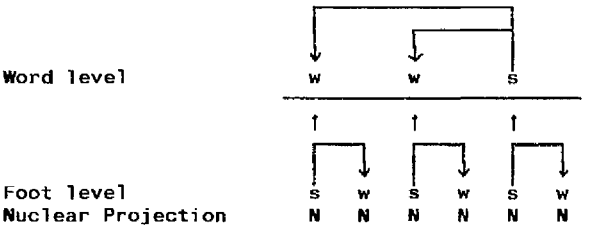
[5]The F here by no means entails the existence of any level of structure different from the actual foot tree level. It is used simply to refer to a foot tree.

(15)



Although the metrical structure in (15) above does not explicitly indicate any government-type of relation between nodes, it reflects the relations of government established among nuclei at the nuclear projection. In the structure above, each strong node governs its weak counterpart at the foot level. At the word tree level, the head of the domain governs each weak node individually. The metrically determined strong node at the word tree level is called the head of the word domain. In this way, the interpretation of metrical structure can be illustrated by the following schema[6,7]:

(16)



[6]Notice that the projections below the nuclear projection have been omitted in the following structures.

[7]In the following diagram, the arrow ↑ refers to the projection of nodes to the upper level of metrical structure.

This sums up the basic theoretical assumptions used in this thesis which are relevant to an understanding of the analysis proposed. For a detailed introduction to GB Phonology, the reader should refer to Kaye (1989, 1990a,b, 1992, to appear), and KLV (1985, 1990).

With respect to the transcription used in this thesis, the current Portuguese orthography is found between double quotation marks, e.g. "carcará". The lexical representation of a form is given between dashes and the IPA is used in those cases, e.g. -mit+o-[8]. The phonetic realization of a form is given in brackets which should not, however, be confused with morphological boundaries, e.g. [pátu]. The context will undoubtedly make the case clear. Translations are given between single quotation marks, as in 'toothpick'.

It is crucial to keep in mind that the data analysed in this thesis come from the Natal dialect, of which I am a native speaker. Natal is a town of around 850,000 inhabitants found in the Northeastern part of Brazil. It is the capital of the State of Rio Grande do Norte.

[8]The morphological parsing is indicated with [ ] or + only where that is relevant. A discussion on this topic is found in chapter 3.

2.1 *Introduction.*

In this chapter, we examine the assignment of stress to verb paradigms in Brazilian Portuguese. We discuss the principles and parameters available in Metrical Theory which are capable of assigning stress to this grammatical category. A full analysis of various aspects of stress assignment will be offered in chapters 2 through to 5, including a full account of the morphological and metrical structures of verbs and the theoretical implications of each type of structure. Several issues are discussed: the presence of analytic domains; the effects of the insertion of clitics; spreading effects; and the processes which take place when identical elements become adjacent. Along with this analysis, a detailed description of the characteristics that distinguish the Natal dialect from all other dialects will be presented.

We will start with the cases where the assignment of stress is more straightforward and will introduce more complex cases as we proceed into the analysis. This means that throughout the presentation some of the verb forms in some conjugations (where the assignment of stress is not straightforward) will be omitted. The analysis will be gradually refined so as to account for each case in further sections of the presentation.

2.2 *The Assignment of Stress to Verb Paradigms.*

According to the traditional literature stress is distinctive in BP (Câmara 1975). Linguists analysing stress in BP give minimal pairs such as [tɛ́rmĭnŏ] "término" '(the) end' vs. [tɛ̀hmínŏ] "termino" '(I) finish', and [pàrá] "Pará" '(State of) Pará' vs. [páră] "para" '(he) stops' as proof of its distinctiveness (primary stress has been indicated by an acute accent). Stress is said to be assigned lexically[9] given the existence of such contrasts.

[9]Cf. Garde (1973).

In this thesis, we propose an alternative analysis whereby stress is assigned by a set of principles and parameters available in the theory. Let us begin by looking at the distribution of stress in BP. BP words can be stressed, in principle, on any one of the last three syllables, as can be seen in (17). For the sake of clarity, we will give examples containing only open syllables:

(17) a. Final Stress

[àkàbărá]	"acabará"	'(He/she) will finish'
[fùžĭrá]	"fugirá"	'(He/she) will run away'
[lɐvărá]	"levará"	'(He/she) will take away'
[tɔmărá]	"tomará"	'(He/she) will take'

b. Penultimate Stress

[àkábă]	"acaba"	'(He/she) finishes'
[fóžĭ]	"foje"	'(He/she) runs away'
[lévă]	"leva"	'(He/she) takes away'
[tómă]	"toma"	'(He/she) takes'

c. Antepenultimate Stress

[àkàbăríămŭs]	"acabaríamos"	'(We) would finish'
[fùžĭríămŭs]	"fugiríamos"	'(We) would run away'
[lɐvávămŭs]	"levávamos"	'(We) used to take away'
[tɔmávămŭs]	"tomávamos"	'(We) used to take'

Of the three patterns of stress present in BP words, penultimate stress ((17)b) is said to be the unmarked case. Nouns and adjectives which display the two other patterns of stress are even considered loans[10]. The assumption that penultimate stress is the unmarked case is based on the fact that most BP words display this pattern of stress. This is also true for words with no apparent (relevant) morphological structure, such as prepositions, conjunctions and adverbs (18).

---

[10]Cf. Mateus (1982), p. 205 (translated by SOS): "We are dealing with [-native] words ... that will have to be marked with a diacritic (for example [+A]) in their lexical entry so that the following rule can apply:

$$\begin{bmatrix} V \\ +A \end{bmatrix} \rightarrow [+stress]/[\_\_\_]C_0\#$$ ";

and (p.206) "Just as the [-native] words, these words (antepenultimately-stressed words, SOS) will also have to be marked lexically with a diacritic (for example [+E] (E stands for "esdrúxula", 'abnormal', SOS)) to mean that a special rule will apply:

$$\begin{bmatrix} V \\ +E \end{bmatrix} \rightarrow [+stress]/[\_\_\_]C_0VCVC_0\#$$ ";

(18) a. Prepositions:

[désdĭ]	"desde"	'since'
[éntrĭ]	"entre"	'between'
[sóbrĭ]	"sobre"	'on'
[pɐ̀rɐ̃tĭ]	"perante"	'before'
[ɐ̃tĭ]	"ante"	'before'

b. Conjunctions:

[kwɐ̃dŭ]	"quando"	'whenever'
[pòhtɐ̃tŭ]	"portanto"	'therefore'
[kɔ̃túdŭ]	"contudo"	'however'
[kómŭ]	"como"	'as'
[ɛ̆bɔ́ră]	"embora"	'although'

c. Adverbs:

[dɨprɛ́să]	"depressa"	'quickly'
[lɔ̃ʒĭ]	"longe"	'far'
[óʒĭ]	"hoje"	'today'
[sɛ́dŭ]	"cedo"	'early'
[nŭ̃kă]	"nunca"	'never'

We will begin this metrical analysis with the most frequent pattern of stress, namely the penultimately-stressed verb forms such as the ones in ((17)b) above. The other patterns of stress will be dealt with in the following sections and in the following chapter.

Before going into the details of the assignment of stress, let us consider the meaning of a set of principles and parameters responsible for the assignment of stress. It has been proposed that phonological systems are of a very simple nature[11]. As the common assumptions state that linguistic systems are of a restricted nature, it should be the case that linguistic theories based on restrictive principles stand a better chance of successfully accounting for phonological processes than theories where devices such as open sets of rules are involved. The credibility of such phonological theories vanishes as the number of possible languages they predict grows. On the other hand, a phonological theory that deals with a limited number of principles and parameters predicts that languages can only vary within very precise limits. If we think about stress systems, the version of Metrical Theory that takes the principles and parameters of Government Phonology into account will predict, among other things, that languages should have a very simple set of principles

[11]Cf. Kaye (1989), chapter 5.

and parameters which assigns stress to the various categories. Since within this framework we specify the direction of government in terms of a parameter which can be of either left or right orientation, it follows that the amphibrach type of ternary feet generated by Halle & Vergnaud (1987), consisting of a head flanked by two weak nodes (⏑ - ⏑), are excluded from this framework. In no way can notions such as extrametricality be expressed in this framework. Notions such as extrametricality, unpredictable and usually said to be lexically marked, are simply incompatible with this restricted phonological theory.

Let us now consider the following data:

(19)	[divĭdídŭ]	"dividido"	'divided'
	[àtŭrádŭ]	"aturado"	'stood'
	[dizăbáfă]	"desabafa"	'(He/she) speaks up'
	[dizăkátă]	"desacata"	'(He/she) disrespects'
	[dizăbrígă]	"desabriga"	'(He/she) makes homeless'

By looking carefully at the examples in (19), we notice that although primary stress is on the penult, secondary stress occurs on the first syllable as well (secondary stress is indicated with a grave accent). These words display the pattern ò ŏ ó ŏ, where o stands for any position at the nuclear projection and ŏ represents an unstressed position at the level of nuclear projection.

Before we discuss the assignment of stress in these words, let us refresh our minds once again about the relevance of the Licensing Principle in this respect. The Licensing Principle is repeated below:

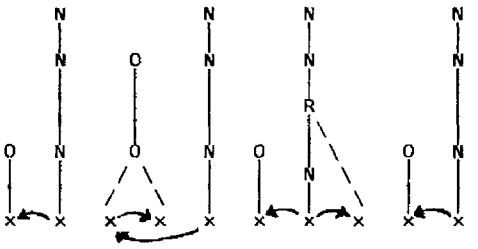
(20) Licensing Principle (Kaye (1990b)

All phonological positions save one must be licensed within a domain. The unlicensed position is the head of the domain.

According to this principle, in a domain, except for the head of the domain all the other skeletal positions must be licensed. What this means is that governees within onsets, nuclei and rhymes are licensed by their head. The head of the onset is licensed by the

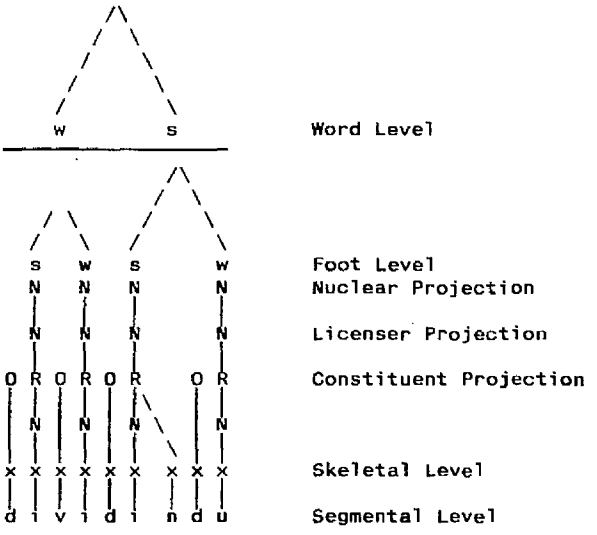
nucleus which follows and which governs it. Governing heads are projected at the licenser projection (Charette (1990)), and after that, nuclear heads are projected to the Nuclear Projection. The licensing at these levels up to the nuclear projection, as just described, is illustrated below:

(21)

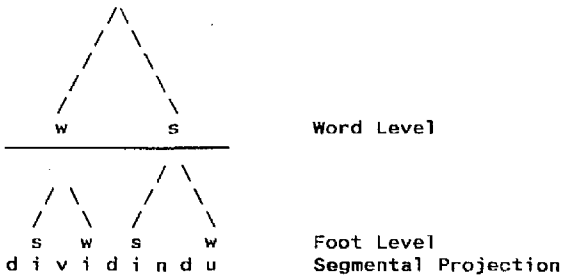


Within the framework of Government-based Phonology, it is at the Nuclear Projection that metrical structure is constructed. This means that only nuclei with phonetic content will be present at this level, since nuclei with no phonetic content (=empty nuclei) have already been licensed at the skeletal level. For typographical reasons, unless the presence of one or more projections proves essential to the discussion in this thesis, metrical structure is constructed on the nuclear heads (segmental level) although we will in reality be referring to the unlicensed nuclei that are projected at the level of nuclear projection. This is illustrated below:

(22) a. Metrical structure where all projections are shown.

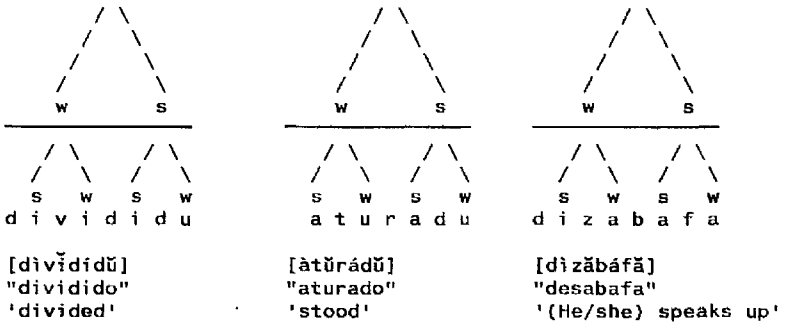


b. Equivalent structure as shown in this thesis.

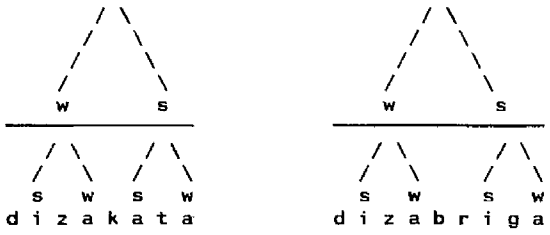


Let us now go back to the examples in (19). The fact that stress is present in alternate syllables suggests the existence of a binary "unit" which in metrical terms is called a Foot. Furthermore, since there is stress on the first[12] and on the third nuclei, feet must be binary and left-headed since the string is scanned in a leftward direction. Following metrical conventions, heads are indicated by s and non-heads by w. Furthermore, since the primary stress is located on the last stressed syllable, we may consider that, at the word level, the head is located on the rightmost node. We propose that the metrical structures of these forms are constructed as shown in (23) below.

(23)



[12]As reported by Major (1985), metrical analyses of stress in BP had already been proposed by Maia (1981). She gives "trees to fit perceived relative prominence of syllables". However, according to Major, she does not attempt to justify these structures.



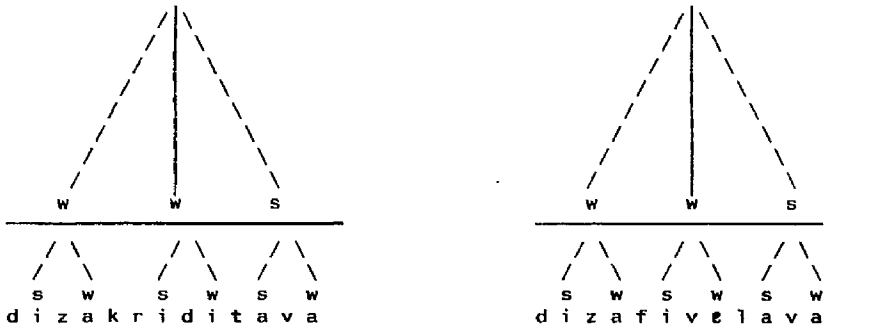
[dìzăkátă] "desacata" '(He/she) disrespects'

[dìzăbrígă] "desabriga" '(He/she) makes homeless'

As shown above, this analysis predicts the right pattern of stress for the words in (19). Translating these observations into metrical settings, we are proposing that feet in BP are binary and left-headed and that, at the word level, one unbounded right-headed tree is constructed. The term "unbounded" is used in Metrical Theory to refer to a metrical constituent whose head is separated from its boundaries by more than one intervening element. In Government terms, this means that the binarity theorem is relaxed at the word level (24).

Let us now consider the following six-syllable words below whose stress pattern can correctly be predicted by the present analysis.

(24)



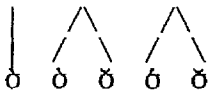
[dìzăkrìdĭtávă] "desacreditava" '(He/she) distrusted'

[dìzăfivɛ̆lávă] "desafivelava" '(He/she) unfastened (belt)'

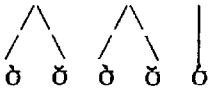
Up to this point we have only considered words with an even number of syllables. However, a question arises as to which edge of the word feet are constructed from. In order to answer this question, we must consider words with an odd number of syllables. If construction is from the right, a degenerate foot will be formed on the left, as in ((25)a). If, on the other hand,

construction is from the left, a degenerate foot will appear on the final syllable, as in ((25)b)[13].

(25) a. Right-to-left



b. Left-to-right



Words with an odd number of syllables in (26) show that the pattern in ((25)a) is correct.

(26)	[pàrávă]	"parava"	'(He/she) stopped'
	[žìrávă]	"girava"	'(He/she) turned'
	[dìzàhŭmávă]	"desarrumava"	'(He/she) made untidy'
	[dìzàpăràfŭzávă]	"desaparafuzava"	'(He/she) unscrewed'

We can therefore conclude that the direction of the construction of feet in BP is right-to-left.

Let us sum up our results to this point.

(27) i. Feet are binary and left-headed.
ii. Foot construction begins at the right edge of the word.
iii. Word trees are right-headed.

We have discussed all parameter settings involved in the construction of metrical constituents in BP based on words that consist of open syllables alone. An appropriate question to ask concerns the cases where heavy syllables are involved. It is well known that stress assignment is sensitive to the branching of

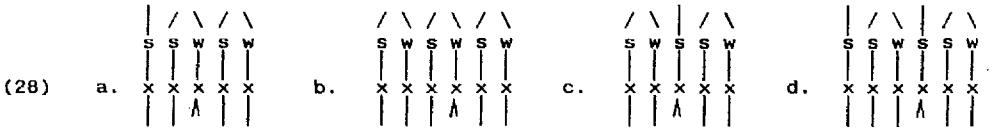
---

[13]The formation of degenerate feet, previously rejected by Liberman & Prince (1977; 256) with the statement "...an isolated [s], an isolated [w] and the configurations [ss] and [ww] are meaningless...", is implemented here. The formation of degenerate feet at the foot level derives from the principle of Metrical Theory which says that every metrical constituent must have a head (Cf. Halle & Vergnaud, 1987, p.9). Major (1985) has given not only phonological but also instrumental evidence for this principle, based on BP data.

nuclear or rhymal constituents in many languages. For example, it has been noted by Liberman & Prince (1977) amongst many others, that stress in Latin is antepenultimate except in the cases where the penult branches: e.g. stress is antepenultimate in [âgríkŏlă] "agrícola" 'agricultural', but it can only be penultimate in [sirkúndŏ] "circúmdo" 'surrounded' because the penult branches. Languages are known to be quantity sensitive or insensitive - quantity sensitive if heavy syllables receive metrical treatment differently from light syllables: otherwise they are quantity insensitive.

Let us now turn to closed syllables in BP. Although metrical structure is constructed on the nuclear projection within the framework of GB Phonology, and nuclear and rhymal complements are absent at this level, we need to check if stress in BP is sensitive to the branching of a nucleus or of a rhyme at the constituent projection. Since foot construction settings are binary and left-headed, we can test whether BP is quantity sensitive by trying to construct a foot where, proceeding from the right, one would have to assign w to a heavy syllable. If BP is quantity sensitive, then a degenerate foot will be formed on the heavy syllables ((28)c, (28)d), otherwise stress will be assigned in the usual way and a degenerate foot will be formed on the first syllable ((28)a, (28)b), in the case of words with an odd number of syllables ((28)a).

(28)



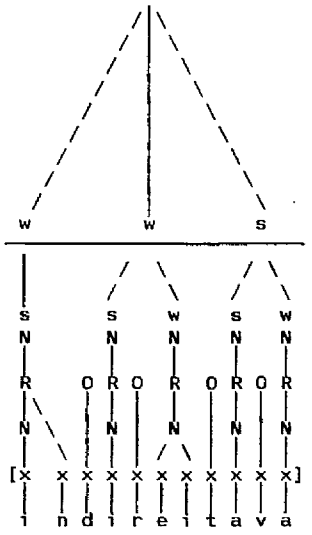
The examples in (29) show that ((28)a, (28)b) are correct. BP is therefore quantity insensitive.

(29)	a.	[ĩdirějtávă]	"endireitava"	'(He/she) tidied up'
	b.	[mànifĕstávă]	"manifestava"	'(He/she) manifested'
	c.	[ĩtrìvĩstávă]	"entrevistava"	'(He/she) interviewed'

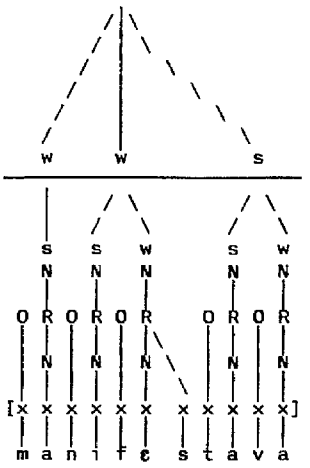
As shown above, stress in BP is insensitive to the branching of

a rhyme or of a nucleus at lower levels of projections. The metrical structure of these forms is constructed as follows[14]:

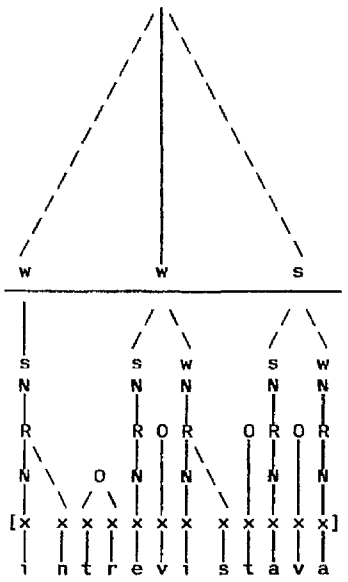
(30)



[ĩdìrẽjtávă]
"endireitava"
'(he/she) tidied up'



[mànìfẽstávă]
"manifestava"
'(he/she) manifested'



[ĩtrìvĭstávă]
"entrevistava"
'(he/she) interviewed'

In this section, we have analysed the third person (henceforth,

[14]The Licenser projection, which is irrelevant to the discussion, has been left out in the following structures.

P3) Indicative Present and the Past Participle (23), and the P3 Imperfect (30). One aspect of the derivation of P2 and P4 in the Indicative Present deserves some attention. Examples of these forms are given below:

(31) Indicative Present[15]

P2	[fɛ̂rĭs]	[bɛ́bĭs]	[mɔ́ʎɐ̆s]
	"feres"	"bebes"	"molhas"
	'(you) hurt'	'(you) drink'	'(you) shower'
P4	[fɨrĭ̃mŭs]	[bɐ̀bɛ̃̀mŭs]	[mɔ̀ʎɐ̃̀mŭs]
	"ferimos"	"bebemos"	"molhamos"
	'(we) hurt'	'(we) drank'	'(we) shower'

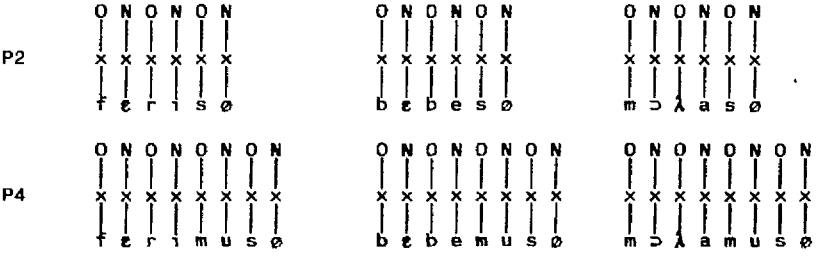
These forms all end in a consonant. According to the Coda Licensing principle (Kaye, 1990b), we can only analyse a word final consonant as being in an onset. This principle requires that every rhymal position be governed by a following onset, as given in (32).

(32) Coda Licensing Principle.

Post-nuclear rhymal positions must be licensed by a following onset.

This principle rules out the possibility of analysing a final consonant as being in a rhymal position, because it is not followed by an onset. And if a final consonant is not in a rhymal position, the only remaining non-nuclear position it can be is the onset. The Licensing principle (Cf. page 16) will predict that this final onset is followed by a nucleus, which in these cases is always empty. The syllabification of these forms is, therefore, as follows:

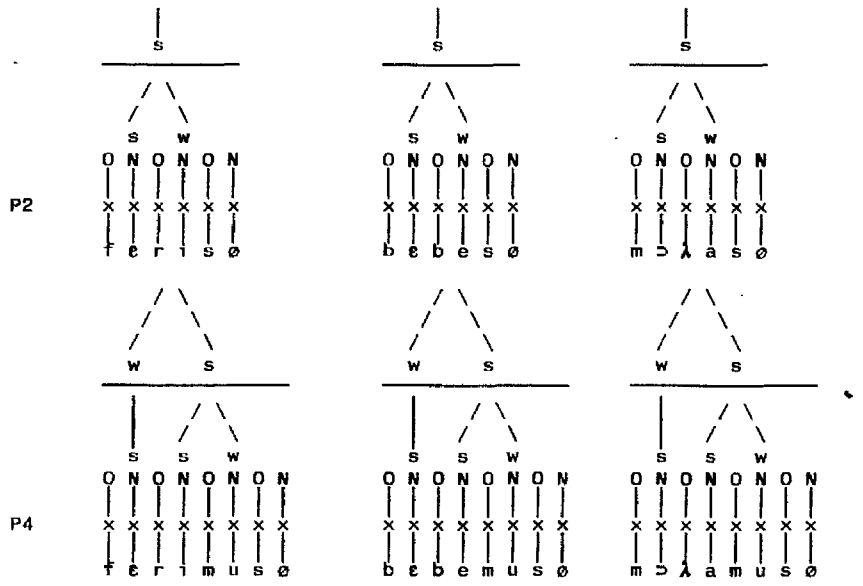
(33) Indicative Present



[15]P1, P5 and P6 are analysed in chapter 5.

According to GB Phonology, nuclei which have no phonetic content are absent from the nuclear projection, the level where stress assignment occurs. This being the case, P2 and P4 should be stressed as follows:

(34) Indicative Present



The pattern of stress predicted by the metrical analysis above is correct, as can be confirmed by the realisation of these forms in (31) (repeated below).

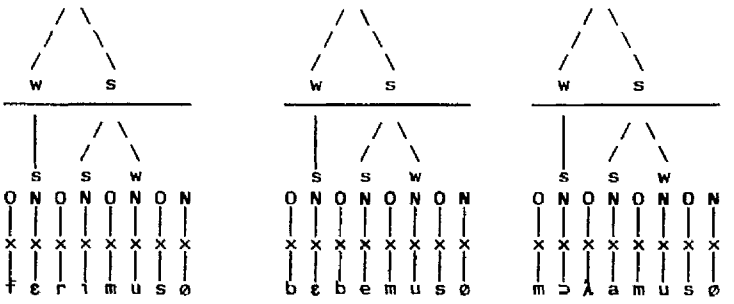
(35) Indicative Present

P2	[fɛ́rĭs]	[bébĭs]	[mɔ́ʎăs]
	"feres"	"bebes"	"molhas"
	'(you) hurt'	'(you) drink'	'(you) shower'
P4	[firĭ̃mŭs]	[bɛbɛ́̃mŭs]	[mɔʎɐ́̃mŭs]
	"ferimos"	"bebemos"	"molhamos"
	'(we) hurt'	'(we) drink'	'(we) shower'

Other finite forms display the same type of structure as shown above. This is indeed the case in the Past P4 and P5 forms, as shown below:

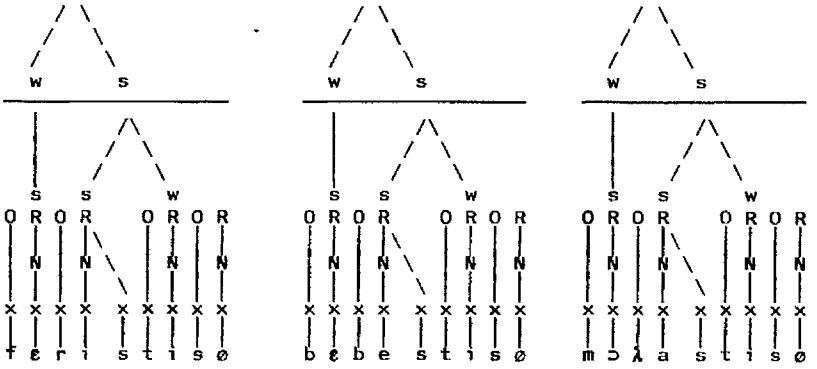
(36) Past

P4



[firĩmũs] [bɛ̀bẽ́mũs] [mɔ̀ʎã́mũs]

P5



[firistĭs] [bèbéstĭs] [mɔ̀ʎástĭs]

In this section, we have analysed the assignment of stress to verb forms. The data analysed have shown that empty nuclei are straightforwardly treated in accordance with theoretical predictions, i.e. they are not taken into account in the construction of metrical structure since they are not present at this level of projection. At this point, let us sum up the metrical structure construction settings proposed for BP once again.

(37) i. Feet are binary and left-headed.
ii. Foot construction begins at the right edge of the word.
iii. Word trees are right-headed.
iv. BP is quantity insensitive.

In the next section, we will be dealing with words demonstrating other stress patterns, together with some related phenomena.

2.3 *Other Finite Forms.*

In the preceding section, we discussed the assignment of stress to the simplest type of verb forms which display the penultimate pattern of stress. In this section we begin the analysis of forms with other patterns of stress.

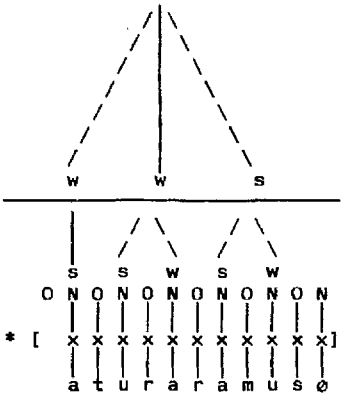
Let us consider the following:

(38) The Pluperfect.

P1 [àtŭráră]
P2 [àtŭrárăs]
P3 [àtŭrárã]
P4 [àtŭrárămŭs]
P5 [àtŭrárẽjs]
P6 [àtŭrárãɯ̃]

The realisations above represent the Pluperfect of the verb "aturar" 'to stand'. Noticeable in this type of structure is that primary stress is columnal, i.e. it always falls on the same nucleus, no matter how many nuclei follow it. According to the analysis of stress proposed so far, primary stress should always fall on the penultimate nucleus with phonetic content. It is obvious that the final nucleus in P2 does not count in the construction of metrical structure, i.e. it is empty. However, according to the metrical settings proposed a form like P4 should be analysed as follows:

(39)



* [àtŭrárámŭs]
[àtŭrárămŭs]

The analysis in (39) is incorrect. Let us therefore investigate the

possible reasons why these forms should be stressed the way they are. In particular, let us see what aspects of the theory could help us in the analysis of these forms.

Let us begin by sketching the organization of domains within this framework. According to GB Phonology, a derivation may consist of any one of four types of domains.

(40) a. Non-Analytic Morphology.

[A B] Domain: AB

b. Analytic Morphology.

i. [[A] [B]] Domains: A, B, AB

ii. 1. *[A [B]]

2. [[A] B] Domains: A, AB

As (40) illustrates, two strings A and B can be concatenated in four different ways. They can be in the same domain ((40)a) or they can be in different phonological domains ((40)b). Although the exact nature of the structure in ((40)a) is still debatable[16], two strings A and B can basically be concatenated in such a way that the juncture is invisible to the phonology. The expectation for this type of derivation ((40)a) is that no phonological process should be sensitive to the environment where the two strings A and B are adjacent. Derivations where the non-analytic type of morphology is observed will behave exactly as unanalysable strings. This is the case in "equal+ity" and "im+possible" in English. In the case of "equal+ity", stress will be assigned to the string [equality] as if no morphological boundary existed and no phonological process will apply at the juncture. According to GB Phonology, "-ity" and "im-" are non-analytic affixes. Another example of a non-analytic derivation is "impossible". In this derivation, the -n- of the prefix in- is in a rhymal position. When the relations of government are established and -n- becomes

[16]According to J. Kaye (PC), it is still debatable whether this type of structure is really decomposable into smaller units or whether it is simply listed as one form with no internal constituency within the lexicon.

governed by -p-, it then assimilates to it. These derivations behave exactly like unanalysable words like 'agenda' and 'cariboo'. 'Equality' and 'impossible' display the structure in ((40)a).

In the analytic type of derivation ((40)b), on the other hand, processes which are sensitive to each environment will apply. Although ((40)b.ii.1) is logically possible, "it does not appear to be attested" (Kaye, to appear). The structure in ((40)b.i) can be illustrated by [[black][board]]. In 'blackboard', both domains receive stress independently and, when the derivation reaches the larger domain, primary stress is assigned to the leftmost domain according to the English Compound Stress Rule (Halle & Vergnaud 1987, p.271), resulting in 'bláckbòard'. Let us consider this derivation in detail. The two domains are [blæk] and [bɔ:d]. Because [blæk] is an independent domain, the final -k- will be syllabified within an onset following the Coda Licensing principle (Kaye, 1990b), discussed earlier (preceding section). This principle precludes the syllabification of -k- within a rhymal position simply because there is no onset following it. The only alternative syllabification for this string is the following:

(41)

Because -k- is in an onset position and according to the Licensing Principle every onset is always followed by a nucleus which licenses it, we assume the existence of a following nucleus (Cf. discussion in the preceding section). Since final empty nuclei are licensed in English, it remains uninterpreted. The same principles apply to the [bɔ:d] domain. The only possible syllabification for this string is the following:

(42)

The next step of the derivation is reached. The boundaries are

deleted and the two strings now become adjacent in the larger [blækøbɔ:dø] domain. The expectations are that nothing happens to the stress assigned in earlier steps of the derivation. This is exactly what happens. In the analytic type of derivation, the individual strings maintain their integrity.

An example of the structure in ((40)b.ii.2) is the derivation of the regular past tense in English. 'Stemmed' is one case where the stem, -stem- in this case, is found in a domain where stress is assigned, and -(e)d is not. What happens basically is that -(e)d- will receive no stress. Because the -m- in -stem- is domain final, it needs to be licensed by the following nucleus, exactly as in the previous case. This implies that the -m- will remain as it is when boundaries are deleted. What we will hear is [stemd].

The structures illustrated in (40) exhaust the possible derivations that languages can display. There is no such thing as post-cyclicity, different layers or levels of phonology or any such concept. It is also important to notice that the non-analytic/analytic distinction is a characteristic of the morphology of affixes in this framework, not of the phonology. Processes take place independently of the nature of the affixes involved and occur at any point in a derivation where their appropriate structure is met.

This discussion brings us back to the forms in (38) (repeated below).

(43) The Pluperfect.

P1 [àtŭráră]
P2 [àtŭrárăs]
P3 [àtŭráră]
P4 [àtŭrárămŭs]
P5 [àtŭrárğjs]
P6 [àtŭrárâ]

From the theoretical point of view illustrated above, the most appropriate analysis for these forms seems to be the one where the person markers are interpreted as analytic suffixes. In order to propose a lexical representation for these forms, let us do their

morphological parsing and then check the most plausible analysis we can give them. Their morphological parsing is given below:

(44) The Pluperfect[17]

P1	a	t	u	r	+	a	+	r a
	STEM					TV		TM
P2	a	t	u	r	+	a	+	r a + s ø
	STEM					TV		TM PM
P3	a	t	u	r	+	a	+	r a
	STEM					TV		TM
P4	a	t	u	r	+	a	+	r a + m u s ø
	STEM					TV		TM PM

As shown above, the Pluperfect consists of the concatenation of a VERB STEM, a THEMATIC VOWEL (TV), a TENSE MARKER (TM), and according to the person, a PERSON MARKER (PM). If we observe the realisations of these forms (43), we notice that, regardless of the presence of a person marker nucleus, stress invariably falls on the thematic vowel. From these examples, we can conclude that the presence of the person markers does not alter the stress in the -aturara- domain. We therefore propose that the reason why these forms are invariably primary stressed on the thematic vowel is because person markers are analytical. This means that the morphological analysis of the Pluperfect from a phonological point of view is the following:

(45) The Pluperfect

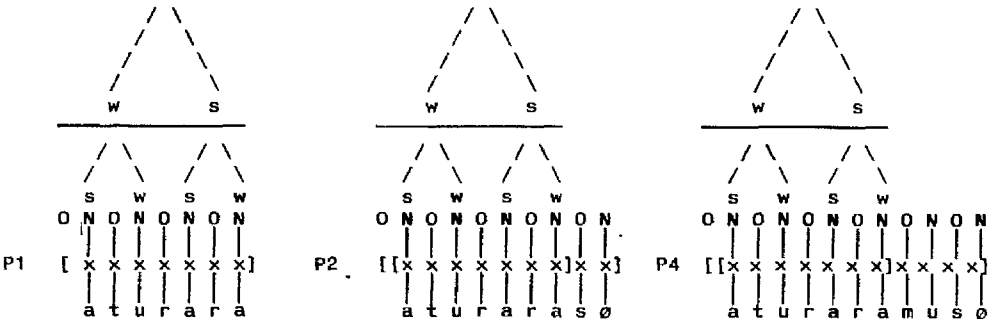
P1	[a t u r a r a]
P2	[[a t u r a r a] s ø]
P3	[a t u r a r a]
P4	[[a t u r a r a] m u s ø]

The square brackets in (45) show that there is an internal domain where the stem -atur-, the thematic vowel -a-, and the tense marker -ra- are present. The person marker, which we propose to interpret as analytic, is found outside the innermost domain and is not surrounded by a set of brackets. It is important to notice that there are no P1 and P3 markers. Stress will be assigned to the innermost domain only, and any string found

[17]P5 and P6, which involve other aspects of the derivation, are only discussed in chapter 5.

outside the innermost domain will not receive stress. (46) below illustrates the metrical analysis of the forms in (45).

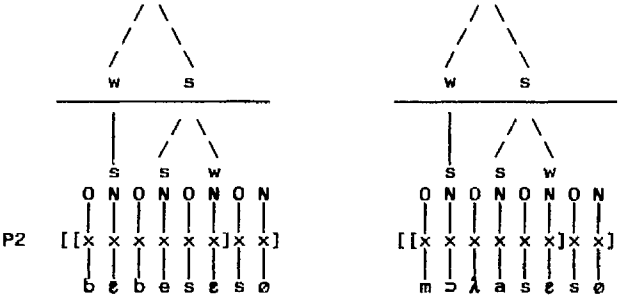
(46) The Pluperfect.



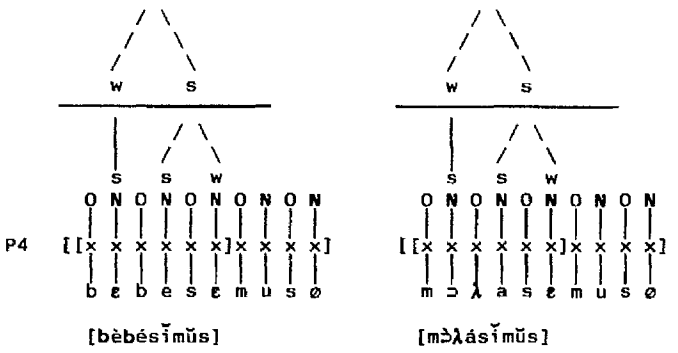
The metrical analysis in (46), where the person markers are interpreted as analytic suffixes, correctly predicts the attested pattern of stress of the forms in (43). In these structures the string found outside the innermost domain, the person markers, receives no stress whatsoever[18].

We propose that this type of analysis also applies to the Subjunctive Imperfect, whose metrical analysis and realisations are found below:

(47) The Subjunctive Imperfect.



[18]The actual reasons why these strings receive no stress at all will be discussed in the next chapter.



This type of structure is common to a few other tenses, such as the Future, the Conditional, and both tenses of the Imperative. However, because the derivation of these tenses involve other aspects which are related to the topics to be discussed in chapters 3 (in the case of the Future and of the Conditional) and 5 (in the case of the Imperative), these tenses will be analysed in chapters 3 and 5, respectively.

We have seen in this chapter how stress is assigned to verb forms in BP. We have proposed that they are stressed with a set of principles and parameters available in the theory.

We have examined a few cases of analytic phonology in verbs as well. We have proposed that person markers are analytic in the Pluperfect and in the Subjunctive Present. We have also proposed that the same type of structure is present in the Future, the Conditional and both Imperative tenses (Negative and Affirmative). By proposing that these suffixes are analytic we automatically account for the stressing of all the data examined.

3.1 *Introduction.*

In the previous chapter we considered how stress is assigned to verb forms in general. We saw that certain forms display the analytic type of structure, and in these cases, the stress patterns assigned in previous stages of the derivation are maintained throughout the derivation.

In this chapter, we will continue the analysis of verb forms whose most plausible phonological analysis involves analytic domains. The future and the conditional are two such cases. Specifically, we will discuss the insertion of clitics in these two forms, since the mesoclitic insertion of clitics - the insertion of clitics in the middle of a syntagm - is only possible in these two tenses.

As a preamble, we start by describing the preferences of the Natal dialect with respect to verb tense choice, clitic usage, and person choice (section 3.2). After that, we will deal with the assignment of stress and related phenomena that take place in these structures when clitics are inserted (section 3.3).

3.2 *Speakers' Preferences with Respect to the Choice of Syntagms and of Clitics.*

As mentioned earlier (cf. Chapter 2), not only are the inflected verbal forms subject to preference in each BP dialect (for example, some dialects use the [tu] "tu" related inflections while others use the [vosé] "você" ones for the second person singular in all tenses), but tenses (the choice of syntagm) are also variable according to speakers' preference. The Future tense (as illustrated in (50)a)) is not used in the Natal dialect nor in most Brazilian dialects. The notion of future is usually expressed with the construction (Personal Pronoun) + inflected verb "ir" 'to go' + Impersonal Infinitive of the main verb.

(48) "(Eu) vou cantar"
P1 go+P1 to sing

'I'll sing or I'm going to sing'

This type of construction (48) makes exactly the same reference to the future as does the English sentence "I will sing tomorrow". Although the situation dictates that the syntagm of Future should be excluded from our analysis because it is so rarely used, there are some very interesting points to be noted about this tense as well as the Conditional. For this reason, in the following section we will analyse the structure of the Future and the Conditional.

Let us now consider some facts related to cliticisation and clitics in Natal.

The clitics which can be used mesoclitically in these two tenses are the following: -u- "him/it", -a- "her/it", -λi- "you (sing.)", -λis- "you (plural)", and the reflexives -mi- (P1), -ti- (P2), -si- (P3 and P6), -nus- (P4), and -vus- (P5).

The possibility of insertion of each of these clitics is of course subject to the type of verb and the grammatical person in question. It is clear that we will only be able to use the reflexives if we are dealing with verbs that admit reflexive pronouns. The remaining clitics can be used with transitive verbs. With respect to the reflexive clitics, their preferred usage in most Brazilian dialects is not the mesoclitic one, but one in which they appear before the verb syntagm in an independent word (the non-cliticised pre-verb syntagm position). This type of structure is illustrated below:

(49) "Eu o enviaria se tivesse tempo"
P1 refl. 'send'+cond.P3 if 'have'+subj.imp.+P1 'time'

'I would send it if I had time'

As mentioned earlier, the syntagm of Future is rarely used. As for the Conditional, the mesoclitic insertion of clitics in this tense is excluded in the Natal dialect. If the verb in question requires any pronominal form whatsoever, its position will be the non-

clitised pre-verb syntagm position. It is important to note that the use of clitics is very restricted in many dialects of Brazilian Portuguese. In some dialects, such as that of Natal, the use of clitics is almost entirely excluded, except in very formal speech.

This concludes the summary of the aspects which are related to the topic of the next section. In the following section, we will deal with Future and Conditional forms, the assignment of stress to these forms, and the insertion of clitics.

3.3 *The Future and the Conditional vs. Cliticisation.*

In the preceding section we discussed some issues relating to the choice of syntagms by Natal speakers with respect to the Future, as well as the form of actual clitics in BP. We now consider the assignment of stress to Future and Conditional strings and the insertion of clitics in these strings.

The Future and the Conditional of verbs "ferir" 'to hurt', "beber" 'to drink' and "molhar" 'to shower' are given below:

(50) a. The Future[19]

P1	[fiɾĭɾéj]	[bèbĕréj]	[mɔ̀lăréj]
P2	[fiɾĭɾájs]	[bɛ̀bɛ̆rájs]	[mɔ̀lărájs]
P3	[fiɾĭɾá]	[bɛ̀bɛ̆rá]	[mɔ̀lărá]
P5	[fiɾĭɾéjs]	[bèbĕréjs]	[mɔ̀lăréjs]
P6	[fiɾĭɾã͡w]	[bɛ̀bɛ̆rã͡w]	[mɔ̀lărã͡w]

b. The Conditional

P1	[fiɾĭɾíă]	[bèbĕríă]	[mɔ̀lăríă]
P2	[fiɾĭɾíăs]	[bèbĕríăs]	[mɔ̀lăríăs]
P3	[fiɾĭɾíă]	[bèbĕríă]	[mɔ̀lăríă]
P4	[fiɾĭɾíămŭs]	[bèbĕríămŭs]	[mɔ̀lăríămŭs]
P5	[fiɾĭɾíĕjs]	[bèbĕríĕjs]	[mɔ̀lăríĕjs]
P6	[fiɾĭɾíã͡w]	[bèbĕríã͡w]	[mɔ̀lăríã͡w]

[19]The actual realisation of the tense marker nucleus -a- as [e] in P1 and the derivation of P4 and P5 are discussed in chapter 5. The derivation of the diphthong in P2 in the Natal dialect is not discussed in this thesis.

The traditional grammar[20] of BP claims that the future and the conditional are derived from the infinitive. As an example, let us consider the following structures:

(51) a. Infinitive

-falarø-	"speak"	'to speak'
-baterø-	"bater"	'to beat'
-banirø-	"banir"	'to ban'

b. Future

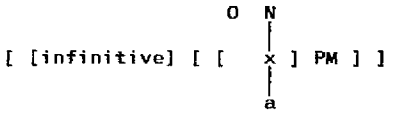
[fàlăréj]	"falarei"	'(I) will speak'
[bàtĕréj]	"baterei"	'(I) will beat'
[bànirréj]	"banirei"	'(I) will ban'

c. Conditional

[fàlăríă]	"falaria"	'(I) would speak'
[bàtĕríă]	"bateria"	'(I) would beat'
[bànirríă]	"baniria"	'(I) would ban'

Even though the assignment of stress to the infinitive is related to the assignment of stress to nouns and adjectives, which is only discussed in chapter 6, for the purpose of the present discussion we will adopt the traditional view that the future and the conditional are indeed derived from the infinitive. For this reason, we will propose that the Future and the Conditional are derived from the following type of structures:

(52) a. The Future



b. The Conditional

Unlike the derivation of the Pluperfect, where only one stress domain exists, we propose that the Future and the Conditional display the analytic type of derivation where two independent domains exist. Since each one of the domains is enclosed within a set of brackets, stress will be assigned to each domain in an

[20]Câmara (1975).

independent fashion. Verbal compounds such as these are possible and have been attested in English. An example of this type of structure is the word "henpecked". Adjectives such as this one derive from the structure shown below:

(53) [[hen] [[peck] ed]]

Let us apply this analysis to attested future and conditional forms in BP.

The application of this analysis for the Future and the Conditional results in the following type of parsing for the forms in (50):

(54) a. The Future

P1	[[fɛrir][[a]i]]	[[bɛber][[a]i]]
P2	[[ferir][[a]sø]]	[[bɛber][[a]sø]]
P3	[[ferir][[a]]]	[[bɛber][[a]]]
P5	[[ferir][[a]isø]]	[[bɛber][[a]isø]]
P6	[[ferir][[a]nø]]	[[bɛber][[a]nø]]

b. The Conditional

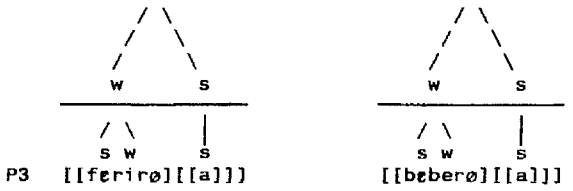
P1	[[ferir][[i a]]]	[[bɛber][[i a]]]
P2	[[ferir][[i a]sø]]	[[bɛber][[i a]sø]]
P3	[[ferir][[i a]]]	[[bɛber][[i a]]]
P4	[[ferir][[i a]musø]]	[[bɛber][[i a]musø]
P5	[[ferir][[i a]isø]]	[[bɛber[[[i a]isø]]
P6	[[ferir][[i a]nø]]	[[bɛber][[i a]nø]]

This means that, in the case of the Future and the Conditional, stress should be assigned to the strings -infinitive- and -tense marker- independently. In a subsequent stage of the derivation, the string -person marker- is incorporated into the -tense marker- string, except for P3 and P1 in the Conditional which have no person marker. Only then do -infinitive- and -tense marker- become members of a single domain.

The bracketing we are proposing, especially with respect to the existence of two independent domains, one for the so-called "infinitive" and another one for the tense marker, gains empirical support from the precise position where clitics can be inserted in these structures. The clitics introduced in the preceding section will, if required, be inserted precisely in the "infinitive" domain, as will be seen from (58) below.

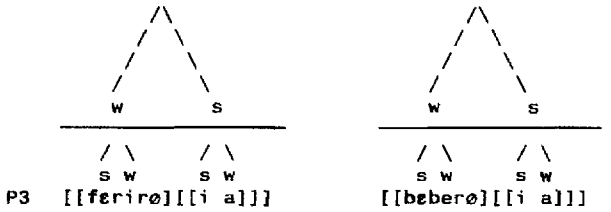
With this bracketing, stress will be assigned to each individual domain, as follows:

(55) a. The Future.



P3 [[fɛrirø][[a]]] [[bɛberø][[a]]]

b. The Conditional.



P3 [[fɛrirø][[i a]]] [[bɛberø][[i a]]]

The derivation in (55) illustrates the assignment of stress as predicted by the analysis. This analysis predicts that there will be primary stress on the tense marker nucleus, in the case of the Future, and on the first nucleus of the tense marker, in the case of the Conditional. It predicts, furthermore, that there will be some degree of stress on the first nucleus of the "infinitive" string, -ɛ- in both cases. This is because, as discussed above, foot construction should not take the empty nucleus after the "infinitive" string into account, since it is not present at the nuclear projection. This analysis correctly predicts the attested pattern of stress for these forms, as given in (50) (repeated below).

(56) a. The Future

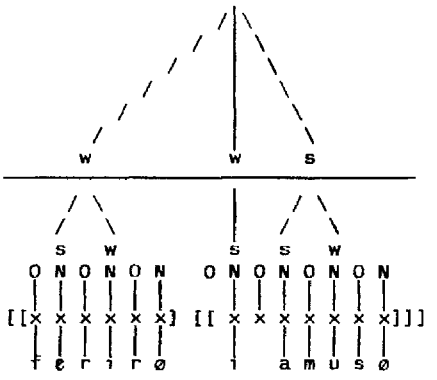
P1	[fìrĭ̌réj]	[bèbĕ̌réj]	[mɔ̀lă̌réj]
P2	[fìrĭ̌rájs]	[bèbĕ̌rájs]	[mɔ̀lă̌rájs]
P3	[fìrĭ̌rá]	[bèbĕ̌rá]	[mɔ̀lă̌rá]
P5	[fìrĭ̌réjs]	[bèbĕ̌réjs]	[mɔ̀lă̌réjs]
P6	[fìrĭ̌ráw̃]	[bèbĕ̌ráw̃]	[mɔ̀lă̌ráw̃]

b. The Conditional

P1	[fìrĭ̌ríă]	[bèbĕ̌ríă]	[mɔ̀lă̌ríă]
P2	[fìrĭ̌ríăs]	[bèbĕ̌ríăs]	[mɔ̀lă̌ríăs]
P3	[fìrĭ̌ríă]	[bèbĕ̌ríă]	[mɔ̀lă̌ríă]
P4	[fìrĭ̌ríămũs]	[bèbĕ̌ríămũs]	[mɔ̀lă̌ríămũs]
P5	[fìrĭ̌ríĕ̃js]	[bèbĕ̌ríĕ̃js]	[mɔ̀lă̌ríĕ̃js]
P6	[fìrĭ̌ríẵw]	[bèbĕ̌ríẵw]	[mɔ̀lă̌ríẵw]

We may consider the possibility that the person markers are incorporated into the same domain as the tense marker, i.e. [[fɛrirø][iamusø]] as an alternative to our proposal, [[fɛrirø][[ia]musø]]. However, this gives a wrong result, as seen in (57).

(57)



* [fìrǐríámŭs]
[fìrǐríǎmŭs]

According to the metrical analysis in (57), primary stress should fall on the second nucleus of the tense marker, -a-, instead of on -i-. This pattern of stress is incorrect. We may conclude, therefore, that the person marker is found outside the rightmost domain.

Let us now discuss the cases where clitics are inserted. As mentioned above, the insertion of clitics is possible in the infinitive domain. In the examples that follow, we will use the clitic -u which, when inserted in the infinitive domain, triggers a phonological change whereby the -r- on the "infinitive" is realised as -l-. (58) below illustrates the realisation of the Future and the Conditional where the clitic -u has been inserted.

(58) a. The Future

P1	[fìrìlŭéǐ]	[bèbèlŭéǐ]	[mɔ̀ʎàlŭéǐ]
P2	[fìrìlŭájs]	[bèbèlŭájs]	[mɔ̀ʎàlŭájs]
P3	[fìrìlŭá]	[bèbèlŭá]	[mɔ̀ʎàlŭá]
P5	[fìrìlŭéjs]	[bèbèlŭéjs]	[mɔ̀ʎàlŭéjs]
P6	[firilŭãw]	[bèbèlŭãw]	[mɔ̀ʎàlŭãw]

b. The Conditional

P1	[firilŭiă]	[bèbèlŭíă]	[mɔ̀ʎàlŭiă]
P2	[firilŭíăs]	[bèbèlŭíăs]	[mɔ̀ʎàlŭíăs]
P3	[firilŭíă]	[bèbèlŭíă]	[mɔ̀ʎàlŭiă]
P4	[firilŭíămŭs]	[bèbèlŭíămŭs]	[mɔ̀ʎàlŭíămŭs]
P5	[firilŭíĕjs]	[bèbèlŭíĕjs]	[mɔ̀ʎàlŭíĕjs]
P6	[firilŭiãw]	[bèbèlŭíãw]	[mɔ̀ʎàlŭíãw]

Here we will ignore the realisation of -r- as -l-, as that would take us far beyond the scope of this chapter.

We have seen that clitics can be inserted in verb paradigms but that their position varies varies according to tense. We have also seen that their use is subject to the speaker's choice. We must now consider what cliticisation really means in phonological terms.

We propose that cliticisation refers to the coalescence of a verb syntagm and a pronoun into one phonological domain. There is a very simple piece of evidence to support this assumption, as shown below (AP stands for Accusative Pronoun):

(59) Cliticisation of a pronoun in the Indicative Present tense[21].

a. [élă] [lávă] [élĭ]
P3(fem) wash+P3PM P3(AP-masc)

"Ela lava ele"
'She washes him'

b. [élă] [[lávă]ŭ]
P3 wash+P3PM+P3(AP-masc)

"Ela lava-o"
'She washes him'

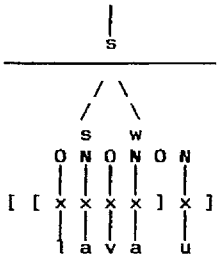
From the two cases illustrated in (59), it is crucial to note that they carry exactly the same meaning. ((59)a) represents the case where the tonic personal pronoun is used as an object. In this case it is stressed in the same way as any other word. This form, which does not contain any clitic, is widely used nowadays in Brazil and is the preferred form in Natal. ((59)b) illustrates the "grammatically correct" form, where the clitic equivalent to the masculine P3 tonic pronoun is used. In this case, it manifests itself as -u and is added to the verb domain.

[21]The bracketing indicates the morphological parsing for the sentence.

What is most important to notice in the case where the clitic is used is the fact that it is unstressed. If it were stressed, then it would necessarily constitute a domain. (59) illustrates cases of cliticisation with the Indicative Present tense. In this tense, clitics follow the person marker.

Cliticisation in the structure in ((59)b) refers not only to the choice of the clitic, but also to the fact that it is adjoined to the innermost domain. In this coalescence, the clitic will not be stressed in the derivation. The organisation of domains in the above structure implies that stress, just like any other phenomenon, will apply to the strings [lava] and [eli] independently in ((59)a), whereas it will be assigned to the innermost domain only in [[lava]u] in ((59)b)[22]. The choice of the clitic over its corresponding pronoun indicates that the former does not constitute a domain on its own. It depends, however, on the verb string it complements. As a clitic, it is found outside the innermost domain in the Indicative Present, and because it occupies this position in the syntagm, it is never stressed. If it were stressed, it would have to constitute a domain. The derivation for the type of form in ((59)b) is given below:

(60)



As shown above, stress is assigned to the innermost domain. Since the clitic is not present in that domain, it does not receive any stress whatsoever.

However, the type of cliticisation seen above refers only to cliticisation in tenses other than the Future and the Conditional.

[22]We will discuss the unstressed status of the clitic below.

In the future and the conditional, the clitics can appear either mesoclitically (58) or in an independent word, in which case they must be positioned before the verb syntagm ((49), repeated below).

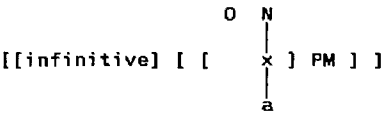
(61) "Eu o enviaria se tivesse tempo"

I would send it if I had time

Let us now concentrate on the insertion of clitics in the middle of Future and Conditional forms. As stated above, the place where clitics can be inserted is not post verb syntagm in these two tenses, in a domain adjoined to the verb syntagm as shown in (60), but between the "infinitive" and the tense markers. The presence of clitics within two parts of a syntagm is, as mentioned above, specific to the Future and the Conditional. We argued that cliticisation refers to the coalescence of a verb syntagm and a pronoun into one phonological domain as shown in ((60)). However, data where the Future and the Conditional are involved reveal that the position is not domain-final in these two tenses, but between the two independent domains (58). This requires further investigation.

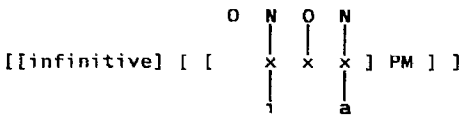
We now concentrate on the structure proposed for the Future and the Conditional. There seems to be an interrelationship between the presence of two individual domains in the Future and the Conditional and the position where clitics can be inserted in these two tenses. No other tenses consist of two domains where stress is assigned independently, and no other tenses allow for the mesoclitic insertion of clitics. This reinforces our proposal that the Future and the Conditional display the structure in (52) (repeated below)[23]:

(62) a. The Future



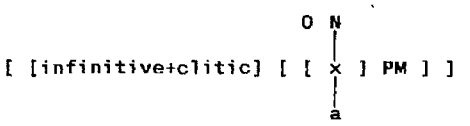
[23]Their full derivations will be found in the appendix.

b. The Conditional

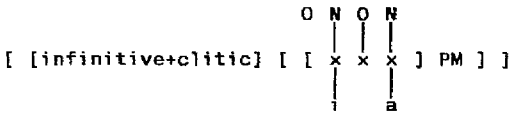


We furthermore propose that clitics, if required, are inserted within the same domain as the "infinitive". The presence of the clitic within the infinitive domain will be responsible for the change in stress in this domain. The structures proposed for the Future and the Conditional where clitics are present are illustrated below:

(63) a. The Future

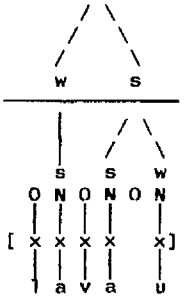


b. The Conditional



It is important to indicate the two positions where clitics can be inserted. We proposed that in the Future and the Conditional clitics appear in the same domain as the so-called "Infinitive" string. On the other hand, in the Indicative Present and all tenses other than the Future and the Conditional, they are found outside the innermost domain. The presence of clitics within the innermost domain in the Indicative Present would give the wrong results, as shown below:

(64)

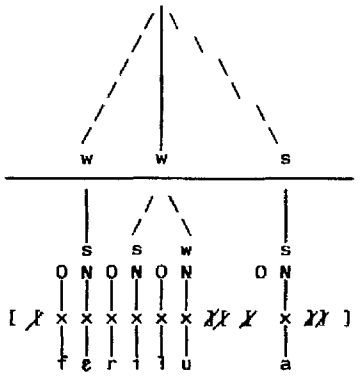


* [làváū]

We can therefore conclude that the interpretation of clitics as

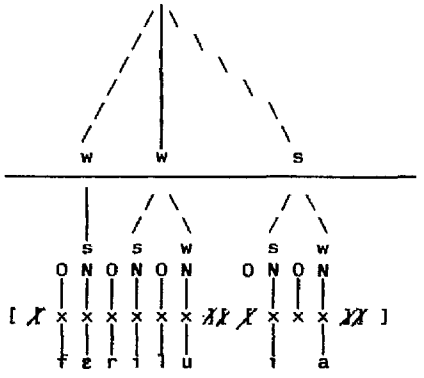
being outside the innermost domain in tenses like the Indicative Present and inside the "infinitive" domain, in the Future and the Conditional, is correct. (65) below represents the derivation of Future and Conditional forms according to the analysis proposed.

(65) a. Future



[fìrilŭá]
"ferir-lo-á"
'(he/she) will hurt him/it'

b. Conditional



[fìrilŭiă]
"ferir-lo-ia"
'(he/she) would hurt him/it'

Following the assignment of stress to the two individual domains, the derivation reaches the larger domain where the word tree is constructed.

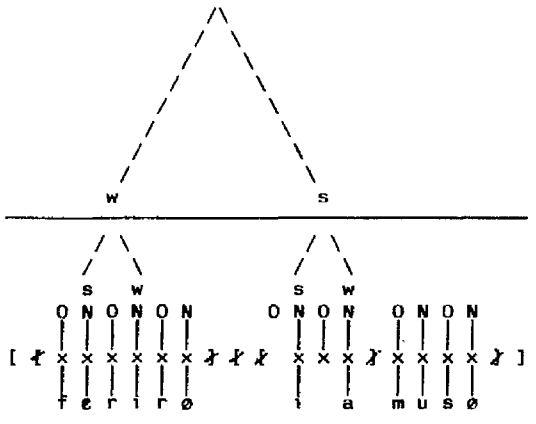
This analysis predicts that with the insertion of the clitic, the only change in stress that occurs is on the "infinitive", in contrast to the forms where clitics are not present. The stress on the tense marker domain remains the same. It also predicts

that primary stress will fall on the rightmost strong node, as in all other cases.

Let us sum up our results so far. We have proposed to analyse the future and the conditional as displaying the analytic type of domain organisation. We proposed that both tenses derive from the infinitive, and that phonologically, both tenses display the parsing [[infinitive][[TM]PM]]. We discussed the cases where clitics are inserted and we proposed that they are added to the -infinitive- domain and that this accounts for the difference in stress between the cases where clitics are absent and when they are present.

Let us now return to the analysis of clitics in the future and the conditional. One fact has not been considered throughout this discussion on clitics and stress assignment in verb forms. We are concerned with the unstressed status of clitics which are found outside the innermost domain in the Indicative Present tense, and nuclei with phonetic content which are appended outside the stress domain on its right. Clitics that are never stressed throughout the derivation arise, as we have seen, in the Indicative Present. Other nuclei with phonetic content which arise in the same circumstances are person markers which are added outside the stress domain in the pluperfect, for example in [[ɛrira]musø]. That is the case of P4 in the Conditional, which is shown below:

(66)



[firĭriămŭs] "feririamos" '(we) would hurt'

The next stage in the derivation outlined above will, of necessity, be one where one nucleus with phonetic content, -u-, is appended domain-finally to a string which has already been assigned stress. We have seen in the attested stress pattern for this form that the final nucleus remains unstressed, as in [fĭrĭríămŭs]. Other than the fact that the destruction of metrical structure of previous stages would violate the Strict Cycle Condition (Kean 1974), we have a very simple mechanism to prevent the destruction of the metrical structure already constructed. Let us simply reconsider the parameter settings proposed for the construction of metrical structure in BP (repeated below):

(67) i. Feet are binary and left-headed.
ii. Foot construction begins at the right edge of the word.
iii. Word trees are right-headed.
iv. BP is quantity insensitive.

In particular, let us recall the directionality of metrical construction proposed in ((67)..ii). (67) dictates that it takes place from right to left, and we expect that if segments are attached to the left of the stressed string in later stages of the derivation, the metrical settings are bound to continue to apply to the string, as long as that does not violate the Strict Cycle Condition. This means that, should we find a case where nuclei with phonetic content are added to the left of a previously stressed string, the new material should receive stress. That would violate neither the Strict Cycle Condition nor the parameter settings proposed for BP.

However, the assignment of stress to the string on the right of the innermost domain needs to be formally ruled out. We therefore propose the following principle:

(68) The No-Backtracking Principle.

The phonological string is scanned from right to left. Throughout the derivation a scanning position must not vary from this direction.

The No-Backtracking Principle, as formulated here, will prevent the stress assigned to the tense marker and all strings to the left of it in previous stages of the derivation from being wiped out.

Without this proviso, feet would be constructed from the person marker leftwards on the string, wiping out all stress to the left of it. Since foot construction has applied to those domains in earlier stages of the derivation, the person marker will not be accessible to the continued application of the metrical structure settings. This analysis correctly predicts the stress pattern of Future and Conditional forms both in the cases where no clitics are inserted and in those cases where clitics are present. A formal explanation for the unstressed status of strings found outside the innermost domain has also been achieved.

4.1 *Introduction.*

In this chapter, we will concentrate on the relations of government expressed through vowel harmony. According to the Licensing principle, unlicensed nuclei present at the nuclear projection contract relations of government. The head of the domain licenses the remaining unlicensed nuclei at that level. One way of expressing these relations of government with each other is through vowel harmony. We have seen in the previous chapter how stress is assigned to verbs. Stress is another instantiation of the governing relations established among nuclei at this level of projection. Following a description of the theoretical issues involved in the representation of nuclear segments according to GB Phonology (chapter 1), in this chapter we analyse the behaviour of mid-vowels in BP.

The data used in this analysis come from verbs. The reason why only verbs are used can be summarised by the following. In order to observe the changes which a specific segment undergoes, we have at our disposal a mechanism that contrasts the unstressed and the primary stressed realisations of particular segments. The possibility of contrasting one unstressed nucleus with its stressed version is only possible in verbs, and, even in verbs, we have to exclude the nuclei which are invariably unstressed (cf. discussion at the beginning of the next section). The possibility of contrasting the primary stressed segment with its unstressed version in other categories is virtually null. However, certain nouns clearly display harmony effects, cf. [âmŏrózŭ] "amoroso" 'loving (masc)' vs. [âmɔ̆rɔ́ză] "amorosa" 'loving (fem)'. In contrast, others clearly deviate from the analysis presented here, cf. [bɛ̆línă] "Belina" 'Belina - model of car'. Vowel harmony effects in other categories are certainly worthy of investigation.

4.2 *Vowel Harmony in the Natal Dialect of BP.*

In this section, the alternations i/e and u/o in pretonic position are examined[24]; hence in all cases we will be dealing with unstressed vowels. More specifically, we will analyse nuclei that may bear primary stress in contrast with cases where they are unstressed (69).

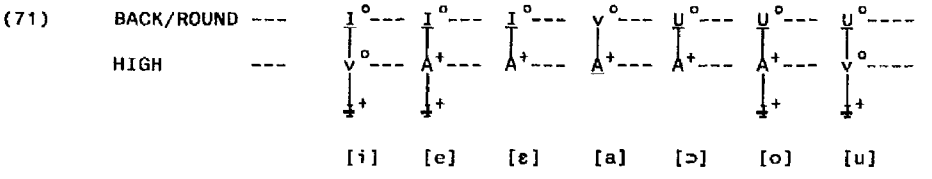
(69)	a.	[kɛ̱brá:vă]	"quebrava"	'(I) used to break'
		[kɛ̱bréj]	"quebrei"	'(I) broke'
		[kɛ̱bri]	"quebre"	'(imperative) break!'
	b.	[fi̱rí̱ă]	"feria"	'(I) used to hurt'
		[fi̱rí:dŭ]	"ferido"	'(past part.) hurt'
		[fɛ̱ri]	"fere"	'(she/he) hurts'
	c.	[kɔ̱lá:vă]	"colava"	'(I) used to glue'
		[kɔ̱léj]	"colei"	'(I) glued'
		[kɔ̱lŭ]	"colo"	'(I) glue'

(69) illustrates cases where a segment varies in complexity. The analysis to be presented here cannot account for cases like the ones in (70) (underlined), because these nuclei will never be primary stressed. It appears that word initial lax mid vowels followed by any nuclei other than lax mid ones (which undergo vowel harmony), do not undergo vowel harmony. These cases cannot be accounted for by our analysis.

(70)	[dɛ̱glŭtiă]	"deglutia"	'(I) used to swallow'
	[ɛ̱làbɔ̱rá:vă]	"elaborava"	'(I) used to elaborate
	[kɔ̱lĭdiă]	"colidia"	'(I) used to collide'
	[kɔ̱lɛ̱sjɔ̱ná:vă]	"colecionava"	'(I) used to collect'

Before we can consider the data, let us remind ourselves of the BP system, which is repeated below:

(71)



The vowel system of BP consists of the seven segments above. There are no lax i's or u's in BP.

---

[24] i/e yield ɛ and u/o yield ɔ in tonic position.

Let us now consider the data. In order to analyse these alternations, let us take simple cases like the ones in (72).

(72)	a.	[fiɾíă]	"feria"	'(I) used to hurt'
		[sitíă]	"sentia"	'(I) used to feel'
	b.	[bèbéŭ]	"bebeu"	'(He/she) drank'
		[pèhdéŭ]	"perdeu"	'(He/she) lost'

By looking carefully at the examples in (72), we notice that the pretonic and the primary stressed nuclei are identical. Let us now compare these structures with those in (73) below, which represent the present tense of the verbs in (69).

(73)	[fɛ́:ɾĭ]	"fere"	'(she/he) hurts'
	[sɛ̃tĭ]	"sente"	'(she/he) feels'
	[bé:bĭ]	"bebe"	'(she/he) drinks'
	[pɛ̃hdi]	"perde"	'(she/he) loses'

By comparing the cases in (72) with those in (73), we notice that the pretonic nuclei of (72), which agreed in quality with the primary stressed nuclei, have now changed. They are still complex, but now primary stressed and lax[25]. We propose that verbs stems in question (73) be lexically represented as shown in (74). We will then have to explain why the $A^{\dagger}$ element is lost and tenseness is added to the pretonic nucleus in the imperfect.

(74) fɛr- sent- bɛb- pɛhd-

We have seen that a segment changes in complexity depending on what segment follows in the primarily stressed position. A nucleus changes from simplex and tense, depending on the segment which follows it, to complex and lax, when it bears primary stress in the present tense. For this reason, in order to propose these lexical representations we must exclude the possibility of the present tense being marked by the presence of the $A^{\dagger}$ element on the primary stressed nucleus. This hypothesis

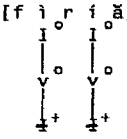
[25]For ease of terminology in this discussion, we reserve the term "complex" here for cases where both the Back/Round and the High lines are active in the representation of a segment (giving mid vowels), and "simplex" for the cases where only one of the lines is active in the representation of a segment. The term "tense" refers to cases where tenseness, the $I^{\dagger}$ element, is present; and "lax" to where it is absent.

could not be true for the simple reason that this would exclude [i] or [u] in primary stressed positions in present tense verbal forms. Examples like these are not only possible but very numerous in BP, as we can see in (75).

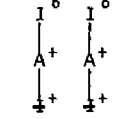
(75)	[diri:ʒi]	"dirige"	'(she/he) drives'
	[tiʒi]	"tinge"	'(she/he) dyes'
	[hedi:mi]	"redime"	'(she/he) redeems'

By now we can be sure that the primary stressed nucleus in some way influences the quality of the pretonic one. We notice that if the primary stressed nucleus (henceforth, N_2) is complex and tense, the pretonic one (henceforth, N_1) is also complex and tense ((72)a); and when the primary stressed nucleus is tense (only), the pretonic will be realised as tense ((72)b). Let us illustrate these effects below (from (72)) by showing how each nucleus is realised:

(76) a. [f i r i ã] "feria" '(I) used to hurt'



b. [b è b é ũ] "bebeu" '(He/she) drank'

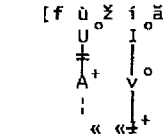


Due to the effects that the primary stressed nucleus imposes on the pretonic nucleus, we propose that these two nuclei hold a relation of government at the level of nuclear projection such that the primary stressed nucleus governs the pretonic nucleus.

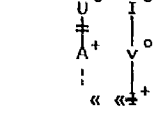
Since we are dealing with pretonic nuclei which become tense in the course of a derivation, it seems important to point out again that there are no lax -i-'s or -u-'s in BP. This means that, whatever the influence that N_2 might have on N_1, if the A^+ line becomes inactive in a given segment in the course of a derivation, I^+ will obligatorily be linked to that segment independently of what segment occupies the primary stressed position.

Let us now consider cases where [ɔ] occupies the N_1 position (77).

(77) [f ù ǯ í ă] "fugia" '(He/she) used to run away'
(Cf. [fɔ́ǯi] "foge" '(she/he) runs away')



[t ù s̆ í ă] "tussia" '(He/she) used to cough'
(Cf. [tɔ́s̆i] "tosse" '(she/he) coughs')



The examples in (77) illustrate another case where N_1 is filled with a lexically complex lax segment (cf. their realisation in the present tense in parentheses) and N_2 is occupied by a less complex governor ([i]). Here again, N_1 is simplified under the presence of a weaker governor (N_2). A^+ is delinked in N_1.

Based on the above facts, we propose the following hypothesis to account for the data both in (72) and in (77):

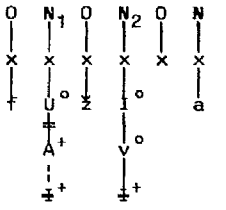
(78) Vowel Harmony in Verbs in the Natal dialect.
(first version)

"N_1 and N_2 are in a governing relation where N_2 is the head. In this governing relation, N_2 is the primary stressed nucleus, the only unlicensed nucleus in the word, which governs all other (licensed) nuclei at the nuclear projection. N_1 is therefore the governee."

From (78) it follows that a complex segment needs to comply with its status as a governee. Theoretically, it must to be no stronger than its governor. We have seen that in all cases N_1 loses an A^+ operator. We therefore argue that, in the particular case of the Natal dialect of BP, governee status entails the loss of A^+ from the governed segment. At this point one must ask why it is that A^+ is lost rather than I° or U°. A possible answer to this question may derive from the fact that A^+ is charmed, whereas I° and U° are charmless. The nucleus in question, being required to be a good governee, needs to lose whatever charmed element it has. However, since lax -i-'s and -u-'s are unpronounceable segments for a BP speaker, $\mathbb{I}^+$ is automatically linked to the weakened segment. Strange though it may seem, the fact that a charmed element is linked, especially in the case where another charmed

element had been deleted due to charm requirements imposed on a governee, somehow does not seem to contradict the charm requirement for the governee[26]. Having ǂ$^{+}$ linked to it, N_1 is now a perfectly pronounceable segment for a Natal speaker. (79) illustrates the process under discussion[27,28].

(79)



[fùžíā] "fugia" '(He/she) used to run away'

As hypothesized in (78), N_1 and N_2 are in a governing relation where N_2 is the head. N_1, being the governee, loses its A^{+} operator (A^{+} is delinked). Since the A^{+} is no longer active, ǂ$^{+}$ links to N_1 in order for I° to become a pronounceable segment in BP. The form [fùžíă] is obtained.

From the data presented so far, one set cannot be accounted for with (78); those in ((72)b) (repeated in (80) below). In these examples, the spreading of ǂ$^{+}$ from N_2 onto N_1 is not predicted (the form [bɛ̀béŭ] should be perfectly pronounceable in Natal). The following discussion will lead us to an account of these cases. Let

---

[26]However, we would like to acknowledge that this problem does not arise in the latest version of the theory. Exactly where our analysis faces difficulty, latest developments in the theory have shown that this is not at all problematic. In its latest version, the number of elements within the GP framework has been decreased by one, i.e. the very ǂ$^{+}$ element no longer exists. Vowel harmony in Natal is now explained by "head alignment" (Kaye PC).

[27]In this analysis, we use | to link elements that are originally (lexically) combined and ¦ to link elements that become combined as a result of spreading.

[28]Since only vowels are at issue in this analysis, the segments in onset positions will all be represented in a simplified way, with IPA symbols.

us consider these data:

(80)	[bèbéŭ]	"bebeu"	'(He/she) drank'
	[pèhdéŭ]	"perdeu"	'(He/she) lost'

We have seen that the realisation of $A^{\dagger}$ operators in a segment occupying the N_1 position is conditional on its presence or absence in the segment occupying the N_2 position. We will now consider the other possibility, namely those cases where $A^{\dagger}$ is the head in the representation of a segment. Let us first examine cases where $A^{\dagger}$ heads are present in the governing position.

(81)	[tìrá:vă]	"tirava"	'(He/she) used to remove' (Cf. [tí:ră] "tira" '(he/she) removes')
	[fèšá:vă]	"fechava"	'(he/she) used to close' (Cf. [fé:ša] "fecha" '(he/she) closes')
	[žɛ̀rá:vă]	"gerava"	'(he/she) used to generate' (Cf. [žéră] "gera" '(he/she) generates')
	[kàlá:vă]	"calava"	'(he/she) used to shut up' (Cf. [kálă] "cala" '(he/she) shuts up')
	[kɔ̀lá:vă]	"colava"	'(he/she) used to glue' (Cf. [kɔ́lă] '(he/she) glues')
	[pùlá:vă]	"pulava"	'(he/she) used to jump' (Cf. [púlă] "pula" '(he/she) jumps')

As the examples above illustrate, a simplex but charmed N_2 can govern almost any segment in the N_1 position, regardless of its complexity. There is only one gap in the distribution of segments in N_1, i.e. the case where it is occupied by -o-. This seems to be an accidental gap. Other than this case, the $A^{\dagger}$ element can govern any nuclear segment in the pretonic position. Let us now consider cases where an $A^{\dagger}$ head occupies the governed N_1 position. This is shown below:

(82)	[pàríŭ]	"pariu"	'(she) gave birth' (Cf. [párĭ] "páre" '(she) gives birth')
	[vàléŭ]	"valeu"	'(it) was worth' (Cf. [válĭ] "vale" '(it) is worth')
	[kàlá:vă]	"calava"	'(he/she) used to shut up' (Cf. [kálă] "cala" '(he/she) shuts up')
	[kàlóŭ]	"calou"	'(he/she) shut up' (Cf. [kàlá:vă] "calava" '(he/she) used to shut up'

[àhú:mă] "arruma" '(he/she) tidies up'[29]

We have seen cases where an A^+ head governs and also cases where it is governed. The realisation of the governee is never influenced by the segment occupying N_2 when -a- is the governing segment. By the same token, the realisation of the A^+ head could not be influenced by any of the possible governing segments since any attempt to weaken [a], such as the change in the head-operator role, would result in an unpronounceable segment for a BP speaker. These facts allow us to conclude that vowel harmony in the Natal dialect is related to A^+ operators only. Specifically, the realisation of an A^+ operator in N_1 will depend on what segment is present in the governing position. This means that our hypothesis should account for three specific types of data: (a) the cases where N_1 contains an A^+ operator but N_2 does not, (b) the cases where N_1 does not contain an A^+ operator but N_2 does, and (c) the cases where both N_1 and N_2 contain A^+ operators. Let us therefore discard (78) and construct a new hypothesis where the A^+ operator role is relevant:

(83) Vowel Harmony in Verbs in the Natal dialect.

(modified version)

"Where N_1 is the pretonic nucleus and N_2 the primary stressed nucleus (head of the domain), the realisation of the governed nucleus (N_1) is directly related to the presence or absence of A^+ operators in both the head and in the pretonic (governed) position:

a. A^+ operators in governed positions (N_1) can only be licensed by A^+ elements in the governing position.

b. $\pm^+$ spreads from N_2 onto N_1 across the A^+ bridge."

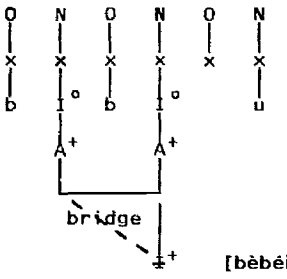
The hypothesis above can not only successfully account for the simplification of N_1 in ((72)a) and (77), but it can also explain the spreading of $\pm^+$ from N_2 onto N_1 in (80).

Before we can analyse each of these derivations, let us define the so-called "A^+ bridge". We use this term to refer to the case where a single element is attached to two adjacent nuclei, as

[29]There are no verb forms displaying this combination, -a- being primarily stressed and -u- following it, in an unstressed nucleus.

shown below:

(84)

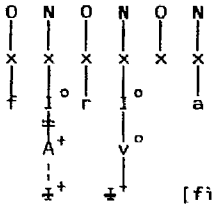


[bèbéŭ] "bebeu" '(he/she) drank'

In the structure above, one single $A^{\dagger}$ element is attached to two adjacent nuclei. In these circumstances, we will say that the $\text{ɬ}^{\dagger}$ element follows the "bridge" and is therefore active in both N_1 and N_2.

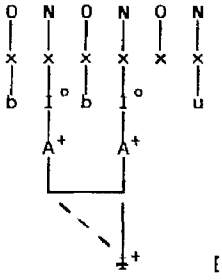
Let us return to the point of our discussion. We have proposed a new hypothesis where the role of the $A^{\dagger}$ element is taken into account. We have also taken into account the spreading of $\text{ɬ}^{\dagger}$ from the head of the governing domain onto its governed position. Let us now consider the data in ((72)a), (77), and (80) (all repeated below).

(85) a.



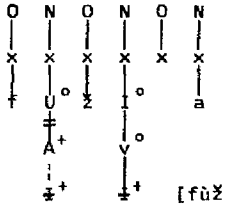
[firíă] "feria" '(he/she) used to hurt'

b.

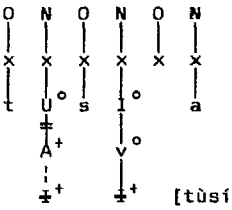


[bèbéŭ] "bebeu" '(he/she) drank'

(86)



[fùžíă] "fugia" '(he/she) used to run away'



[tùsíă] "tussia" '(he/she) used to cough'

Let us now analyse each of the derivations of the forms in (85) and (86). These are cases where (83) applies. In the first example ((85)a), the $A^{\dagger}$ operator in N_1 is not licensed, therefore it delinks. Since the $A^{\dagger}$ line becomes inactive, the $ɬ^{\dagger}$ is automatically linked to N_1 in order for it to become a pronounceable segment for a Natal speaker. That is also the case with -fùžír- and -tùsír- in (86). In ((85)b), both N_1 and N_2 contain $A^{\dagger}$ operators. The $A^{\dagger}$ element in N_1 is therefore licensed to remain there. In that case $ɬ^{\dagger}$ spreads from N_2 onto N_1 across the $A^{\dagger}$ bridge, according to ((83)b). Our hypothesis now accounts for all the data presented.

(83), as it is, precludes us from interpreting the realisation of N_1 as tense in cases where an unlicensed $A^{\dagger}$ operator has been deleted. In these cases, we have no choice but to attribute the presence of $ɬ^{\dagger}$ to the fact that there are no lax -i-'s or -u-'s in BP. N_1 and N_2 do not share $A^{\dagger}$ operators and the spreading of $ɬ^{\dagger}$ is limited to those cases where there is an $A^{\dagger}$ bridge (Cf. ((83)b). This is the case with the realisation of N_1 in ((85)b).

Another set of data can be successfully accounted for with the hypothesis in (83); the cases where the thematic vowel -e- becomes nasalised. This is true in the case of -kɔ̀lɛ́mŭs-, which is illustrated below:

(87) [kɔ̀lɛ́mŭs] "colhemos" '(we) pick up'
(Cf. [kòléŭ] "colheu" '(he/she) picked up')

As illustrated above, the thematic vowel -e-, which is lexically tense, will be unable to spread $ɬ^{\dagger}$ onto the pretonic nucleus simply because it has become nasalised by the presence of a nasal consonant in the following onset. Once it has become nasalised, it has lost its tenseness. Therefore, it no longer has the $ɬ^{\dagger}$

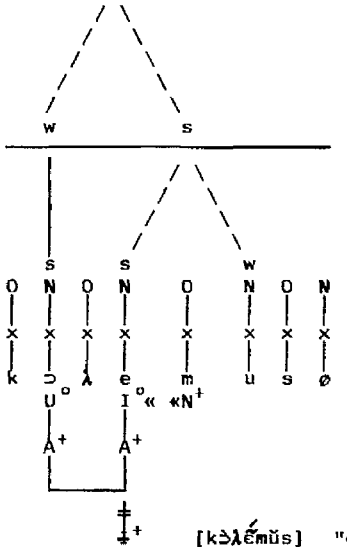
element to spread. This aspect is not limited to BP. It follows from Charm Theory. According to the Charm Theory, elements with like charm cannot combine. Let us examine this aspect in more detail.

According to the lexical representation proposed, the thematic vowel -e- is lexically tense. Its representation according to Government Phonology is given below:

(88) $(\mathbf{I}^{+}.(A^{+}.I^{\circ})^{\circ})^{+}$

As illustrated above, the expression above is positively charmed. The $\mathbf{I}^{+}$ element, which is positively charmed, combines with the $(A^{+}.I^{\circ})^{\circ}$ expression and yields a positively charmed expression (88), or [e]. Nasalisation, on the other hand, refers to the spreading of the N^{+} element, which is positively charmed, onto a segment. Charm Theory dictates that elements with like charm cannot combine, but nasality from an onset which is sandwiched in the rightmost foot obligatorily spreads onto the primary stressed nucleus in the Natal dialect (shown below).

(89)



[kɔ̀lɐ́mũs] "colhemos" '(we) pick up (pres.)'

As shown above, nasality spreads from the sandwiched nasal consonant onto the preceding primary stressed nucleus. Because there is a charm clash between the nasal element and the $\mathbf{I}^{+}$ (both are positively charmed) and nasalisation is obligatory in these

circumstances in Natal, the $\mathbf{I}^{+}$ element delinks. It is therefore inactive, and thus unable to spread onto the governed N_1 position. What we hear is [kɔ̀λɛ̃̀mŭs].

Let us now bring into this discussion other types of data where the derivation is less straightforward. According to (83), an A^{+} which is unlicensed is obligatorily deleted, but there is no obvious reason why a licensed A^{+} should be deleted. Nevertheless, such cases abound. We will consider these below. Throughout this analysis we have considered the behaviour of lax mid vowels that can be primary stressed, since they are the only ones that can undergo simplification. Our main interest has been to see what processes they undergo according to the primary stressed nucleus that follows them in the N_2 position. Since we already know that A^{+} heads play no role in these cases, we list below all the remaining relevant combinations, where the first segment represents N_1, the pretonic nucleus, and the second segment represents N_2, the head of the domain:

(90) a. i. ɛ...i
ii. ɛ...e
iii. ɛ...ɛ
iv. ɛ...ɔ
v. ɛ...o
vi. ɛ...u

b. i. ɔ...i
ii. ɔ...e
iii. ɔ...ɛ
iv. ɔ...ɔ
v. ɔ...o
vi. ɔ...u

According to the hypothesis in (83), the respective realisations for the combinations above should be as follows:

(91) a. i. i...i
ii. e...e
iii. ɛ...ɛ
iv. ɛ...ɔ
v. e...o
vi. i...u

b. i. u...i
ii. o...e
iii. ɔ...ɛ
iv. ɔ...ɔ
v. o...o
vi. u...u

In most cases, the analysis presented makes the correct predictions. In certain cases, however, different results are attested. We now discuss the attested realisations.

Let us consider the first combination ((90)a.i). The combination ɛ...i should yield only i...i, according to (83). However, the following forms are attested:

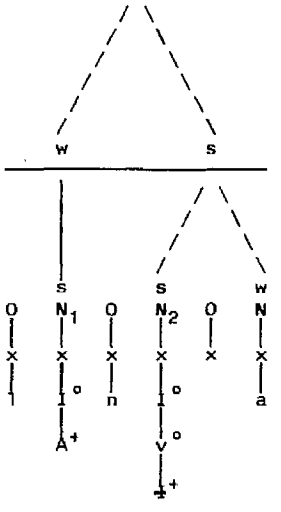
(92) a. ɛ...i Two examples: [lɛ̃níă] "lenia" '(he) used to lenite', and [frɛ̃míă] "fremia" 'to roar, to tremble'

b. e...i Two examples: [kɔ̃nvěhžíă] "convergia" 'to converge', and [divěhžíă] "divergia" 'to diverge'

c. i...i All other cases, ex: [fìríă] "feria" '(he/she) used to hurt'

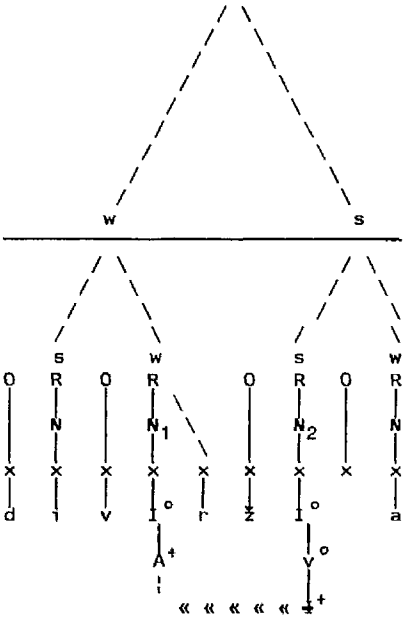
As shown in ((92)a), in two cases unlicensed $A^{\dagger}$ operators are not delinked when followed by a primary stressed -i-. In two other cases, not only do unlicensed $A^{\dagger}$ operators remain linked, but $\downarrow^{\dagger}$ spreads from the head of the governing domain as well. These two processes are, respectively, illustrated below:

(93) a.



[lɛ̀níă]

b.



[divěržíă]

As illustrated above, the $A^{\dagger}$ operator in N_1 is not licensed by any $A^{\dagger}$ operator in N_2. It should therefore be deleted. The only predicted realisation for N_1 should be [i]. These cases, even

though speakers show some degree of uncertainty as to how to pronounce them, constitute exceptions to (83). This uncertainty with respect to the pronunciation of these forms seems to be linked to the relatively low frequency with which these words are used. The remainder of the attested forms are predicted by our analysis, i.e. N_1 is realised [i].

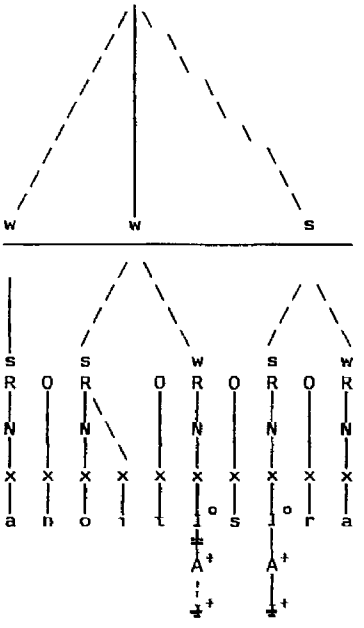
The next set of realisations is related to the combination ɛ...e ((90)a.ii). According to (83) the only result from this combination should be e...e. However, the following types of realisations are attested:

(94) a. e...e ex: [bèbéŭ] "bebeu" '(he/she) drank'

b. i...e ex: most verbs ending in "-ecer", e.g. [ànɔ̀jtĭséră] "anoitecera" '(it) had become night'

As illustrated above, most verbs ending in the "-ecer" '-ise-' suffix undergo a process which is exactly the reverse of the one just seen in the preceding paragraph. In this combination, quite unexpectedly, an A^+ operator which is licensed (by another A^+ element in the head of the domain) is delinked. This process is illustrated below:

(95)



[ànɔ̀jtĭséră] "anoitecera" '(it) had become night'

Although the reason why this licensed $A^{\dagger}$ operator in the pretonic is delinked still needs some clarification, this phenomenon seems to have at least one correlate: the frequency with which the verb is used. For example, -anoitɛsera- "anoitecera" 'it had become night' and -falɛsera- "falecera" '(he/she) had passed away', which are used very frequently, are realised [ànòjtĭséră] and [fàlĭséră], whereas -ɛnaltɛsera- "enaltecera" '(he/she) had exalted' and -mɛrɛsera- "merecera" '(he/she) had deserved', which are less frequently used, are realised [ɛ́nàɯtĕséră] and [mèrĕséră], respectively.

The next combination, ((90)a.iii), and its counterpart where -ɔ- is linked to the pretonic nucleus ((90)b.iii), is related to those cases where a thematic vowel is followed by a nasal consonant. These cases belong to the categories discussed on page 58. Further to that discussion, these cases are not, therefore, considered exceptions to (83).

There are cases where there are no paradigms to illustrate the expected outcome. This is the case of the expected realisations of ((90)a.iv), ((90)a.vi), ((90)b.iv), and of ((90)b.vi). These four gaps are due to the absence of -o- and -u- thematic vowels.

The ɛ...o ((90)a.v) combination displays only its predicted outcome, as illustrated below:

(96) i. e...o ex: [žèlóŭ] '(It) became cold'

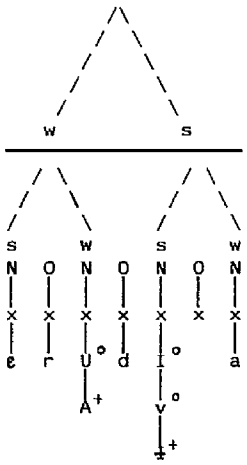
Let us now consider the attested realisations where the pretonic nucleus is -ɔ-. The first case relates to the presence of -i- in N_2. According to the vowel harmony analysis summarised in (83), this combination should only yield u...i. However, in five cases the unlicensed $A^{\dagger}$ operator in the pretonic does not delink. The relevant data are given below:

(97) a. ɔ...i Five cases: [isplɔ̃diă] "explodia" '(it) used to blow up', [istɔ̃hkiă] "estorquia" '(he/she) used to extort', [ɛ̀rɔ̃diă] "erodia" '(it) used to erode', [ɛ̀klɔ̃diă] "eclodia" '(it) used to blow up', and [ĩ̃plɔ̃díă] "implodia" '(it) used to implode'

b. u...i All other cases, ex: [bùlíă] "bolia" '(he/she) used to browse'

The process under discussion is illustrated below:

(98) a.



[ɛ̀rɔ̃díă] "erodia" '(it) used to erode'

In (98), the $A^{\dagger}$ operator is not licensed because there is no $A^{\dagger}$ element in the primary stressed nucleus. It should, therefore, delink. The five verbs in ((97)a) constitute exceptions to (83). All other cases display the outcome which is predicted by our analysis. It is important to notice that the fact that these verbs begin with a nucleus plays no apparent role in this respect, since other nucleus initial verbs like -ɛlɛžera- '(he/she) had elected' and -ingɔlira- '(he/she) had swallowed', which surface [èlɛ̌žéră] and [ĩgũlíră], respectively, do undergo vowel harmony.

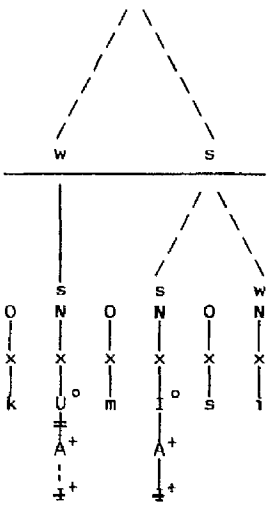
The next combination, ɔ...e ((90)b.ii), should yield only o...e as its outcome. The data given below show that there are two exceptions to the predicted result:

(99) a. u...e Two cases: [kùmésĩ] "comesse" '(if he/she) ate', and [šùvésĩ] "chovesse" '(if it) rained'

b. o...e All other cases, ex: [kòléĩ] "colei" '(I) glued'

In these cases, a licensed $A^{\dagger}$ element in the pretonic unexpectedly delinks exactly like the exceptions given in (94). This process is illustrated below:

(100) a.



[kùmésĭ] "comesse" '(if he/she) ate'

The final combination to be considered relates to the ɔ...o ((90)b.v) combination. The only expected outcome to this sequence should be o...o. The past tense of three verbs yields forms where again a licensed A^{+} element is delinked. The data under discussion are given below:

(101) a. u...o Three cases: [tùmóŭ] "tomou" '(he/she) drank', [bùtóŭ] "botou" '(he/she) put', and [àmmŭsóŭ] "almoçou"'(he/she) had lunch'

b. o...o All other cases, ex: [kòlóŭ] "colou" '(he) glued'

The cases in ((101)a) are related to one set of exceptions which display -a- (an A^{+} head) in N_2. The forms [tùmóu], [bùtóu] and [âlmŭsóŭ] represent the P3 past form of this unmentioned set of exceptions. In the imperfect, -tɔmava- [tùmávă] "tomava" '(he/she) used to take', -bɔtava- [bùtávă] "botava" '(he/she) used to put', and -almɔsava- [àmmŭsávă] "almoçava" '(he/she) used to have lunch', the A^{+} operator in N_1 is also deleted. According to our discussion on page 54, the A^{+} head in N_2 should be able to license any segment in N_1. Yet, in these three cases, the A^{+} operator in N_1 is deleted. However, due to the low number of exceptions where an A^{+} head is present, we will attribute this phenomenon to an idiosyncrasy of these verbs.

The exceptions considered from (92) onwards summarise all

exceptions to the analysis of vowel harmony in verbs proposed in this chapter. We will now speculate as to a possible explanation for these exceptions.

Starting with the i...e outcome in ((94)b), the fact that these cases are on the increase and that there are two cases where a licensed $A^{\dagger}$ element delinks in the pretonic suggests the potential of the dialect for incorporating the delinking of licensed $A^{\dagger}$ operators, as described above, as part of the harmony process.

There are, however, other aspects of these exceptions which seem to play a role in the processes described in (92): the presence of I° as the head of the primary stressed nucleus[30]. (102) below summarises all cases of exceptions seen in this chapter:

(102) Exceptions to (83):

a. ɛ...i	→ ɛ...i	Two examples.
b. ɛ...i	→ e...i	Two examples.
c. ɛ...e	→ i...e	Most verbs with "-ecer" '-ise'.
d. ɔ...i	→ ɔ...i	Five examples.
e. ɔ...e	→ u...e	Two examples.
f. ɔ...a	→ u...a	Three examples.
g. ɔ...o	→ u...o	Three examples.

Nearly all exceptions are cases where the I° element is the head of the primary stressed nucleus. We will conclude that the simplification of the pretonic (delinking of licensed $A^{\dagger}$ operators) is motivated by that fact.

In these cases, regardless of what segment occupies the N_1 position, the N_2 position is always occupied by either -e- or -i-. According to the segmental representation proposed in Government Phonology, e/i should behave exactly like o/u, since the first segment of each pair is composed of a neutral element, I° or U°, as the head of the expression, and $A^{\dagger}$ and $\text{I}^{\dagger}$ as operators in each

[30]The final two cases, which are interrelated, are considered to be an idiosyncrasy of the verbs in question (Cf. discussion on page 64).

pair. It has been noted[31] that U° and I° display dissimilar behaviour in Japanese. When U° is not governing, for example, it delinks. When, on the other hand, it is governing, it surfaces with no lip rounding or lip compression at all. The behaviour of I°, on the other hand, does not vary according to its governing status. What the data from BP suggest is that whenever the head of a harmony domain is a segment whose head is I° and it governs a complex segment occupying the N_1 (pretonic) position in a verb paradigm, it lacks the power to do so. As a result, the segment in N_1 is simplified, and once again in BP A^+ is delinked. In contrast with Japanese, I° seems to be the weaker element in BP, and it is not powerful enough to govern a complex segment. However, in BP, when U° is the head of the expression occupying the N_2 (primary stressed) position only three exceptions exist. U° does not trigger the delinking of either I° or U° from the expression occupying the N^1 position.

In this chapter we have analysed the behaviour of mid vowels in pretonic position in verb paradigms. Specifically, the data examined have shown that harmony in the Natal dialect of BP is related to the licensed/unlicensed status of A^+ operators in governed positions (pretonic nuclei). We have concluded that A^+ operators in N_1 (the pretonic nucleus) which are not licensed by another A^+ in N_2 (the primary stressed nucleus) do not surface.

We have seen that the relation between vowel harmony and stress is only indirect. They both reflect the relations of government among nuclei. The head of the domain governs all other unlicensed nuclei.

---

[31]Cf. S. Yoshida (1991).

5.1 *Introduction.*

In this chapter we will discuss some verb forms where the derivations are less straightforward.

Traditional grammars of Portuguese say little about verb morphology or paradigm derivation and even less about vowel harmony in verbs. It is understandable that vowel harmony has not attracted much attention, simply because very few dialects display it[32]. With respect to the morphological parsing of verbs in general, this topic is discussed very superficially even in the most recent works on this language[33].

Some person markers or whole tense paradigms display a somewhat less than straightforward derivation. This is the case in paradigms which are characterised by adjacent nuclei in the morphology such as P1 Indicative Present, as in [beb+e+u]; all forms of the Subjunctive Present; P1 and P3 of the Simple Past Indicative; and the Affirmative and the Negative Imperatives. In these cases, harmonized nuclei surface with primary stress. In this chapter we will deal with these cases with more problematic derivations, such as the derivation of P5 in most tenses and of P4 in the Indicative Future. We will start with the derivation of P1 in the Present Indicative together with similar derivations.

5.2 *Indicative Present P1 and Similar Derivations.*

In all cases of vowel harmony examined so far (chapter 4), the nuclei which undergo vowel harmony are either unstressed (=weak member of a foot) or secondarily stressed, but never primarily stressed, since vowel harmony is always triggered by the primary

---

[32]At least the type of vowel harmony discussed in chapter 4.

[33]Cf. Mateus et al (1989), and Parkinson (1988).

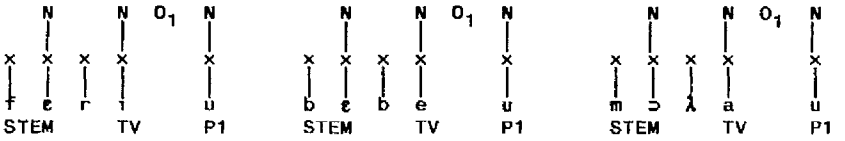
stressed nucleus. However, a few forms display nuclei which, according to the analysis proposed in chapter 4, seem to have undergone vowel harmony and yet still bear primary stress. This is the case of P1 Indicative Present, the Subjunctive Present and both Imperative tenses (affirmative and negative). Examples of these forms are given below:

(103) *Indicative Present.*

P1 [fírŭ] 'I hurt' [bébŭ] 'I drink' [mɔ́λŭ] "I shower"

According to the vowel harmony analysis proposed in chapter 4, the unlicensed $A^{\dagger}$ in the final stem nucleus of the verb -fɛr- 'hurt' should be delinked and $\mp^{\dagger}$ should be added to it, yielding [i], when the primary stressed thematic vowel -i- follows it. The vowel of -bɛb- 'drink' should become tense ([e]) when the primary stressed thematic vowel -e- follows it. On the other hand, the stem nucleus in -mɔλ- 'shower' should remain unchanged when it is followed by the primary stressed thematic vowel -a-. Even if we ignore the place where primary stress falls and the actual realization of the thematic vowels in the paradigms in (103) (where they are in fact inaudible), the actual realisation of the stem final nuclei in these cases ([i], [e], and [ɔ], respectively) suggests the presence of their respective thematic vowels in their lexical representation. The actual nature of the derivation in each case will be investigated in depth as we proceed with the analysis of these forms. Also, because all three forms surface with a final [u], we propose that the first person marker in this tense is -u-. This leads us to propose the lexical representation for the P1 Indicative Present as being:

(104)



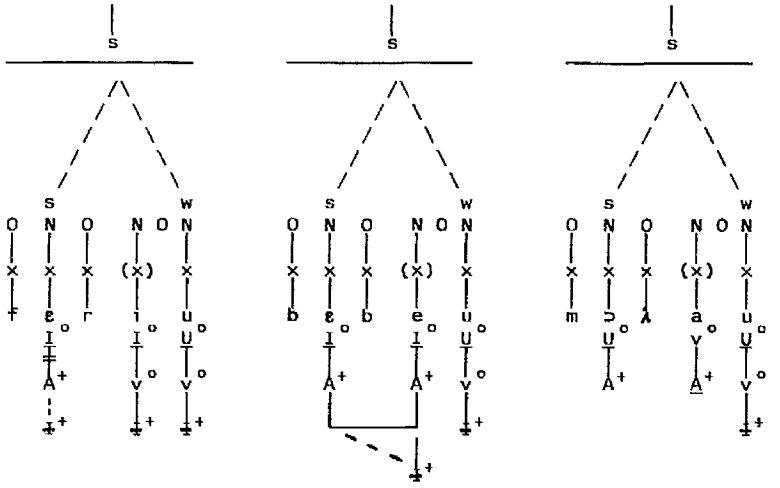
In each of the above cases, the stem is followed by the thematic vowel and the person marker, -u-. This type of structure is new to our analysis in one important respect: we are proposing that it has a sequence of two adjacent nuclei with no intervening onset

point. We propose that this aspect of the representation, namely the succession of two nuclei with no intervening onset position, is responsible for the nature of the derivation. For these structures, we propose the following series of events alongside vowel harmony. First of all, the adjacency of the two nuclei triggers the loss of the leftmost skeletal point by application of the Obligatory Contour Principle (indicated by the parentheses in the following derivations). Although the thematic vowel has lost its skeletal point and is not primarily stressed, it still harmonises the preceding nucleus, precisely as proposed in chapter 4. Because it has no skeletal point, it is not projected for metrical structure and, for the same reason, is inaudible. Let us refresh our minds of the metrical structure construction settings proposed for Brazilian Portuguese. They are repeated below:

(105) i. Feet are binary and left-headed.
ii. Foot construction begins at the right edge of the word.
iii. Word trees are right-headed.
iv. BP is quantity insensitive.

When these metrical settings apply, primary stress is assigned to the harmonised nucleus. This derivation is illustrated below[34].

(106)



As a result of vowel harmony, the pretonic nuclei in (106) should be realised [i], [e] and [ɔ], respectively. The processes just

[34]In the derivations that follow, the parentheses represent the loss of the skeletal position that occurs as described.

described are confirmed by their realisations in (103) (repeated below).

(107) *Indicative Present.*

P1 [firŭ] 'I hurt' [bébŭ] 'I drink' [mɬλŭ] "I shower"

The analysis proposed for P1 Indicative Present also applies to other paradigms which, as we shall see, similarly display a succession of nuclei in their morphology. Although only P1, P4 and P6 of the Subjunctive Present are given (108), there are other paradigms whose lexical representations and derivations are similar to them, but for typographical reasons they will only be given in the appendix which is found at the end of this thesis.

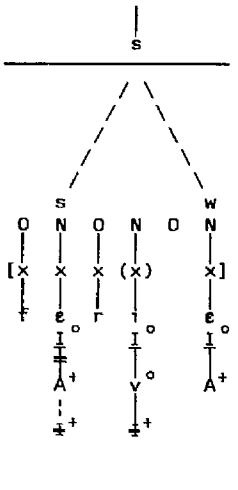
Since many of the derivations of both the Affirmative and the Negative Imperatives present the same type of lexical representation and undergo exactly the same series of events as the Subjunctive Present, we will only illustrate the Subjunctive Present here, and include all the other tenses in the appendix[35,36].

[35]P5, which is excluded in all tenses and conjugations in (108), will be discussed below. In the following examples, P2 and P3 have been omitted because they display the same type of derivation as P1.

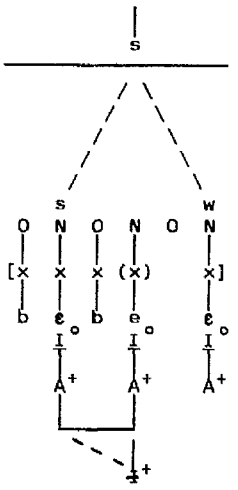
[36]The realisation of the tense marker in (108) is discussed below.

(108) a. Subjunctive Present

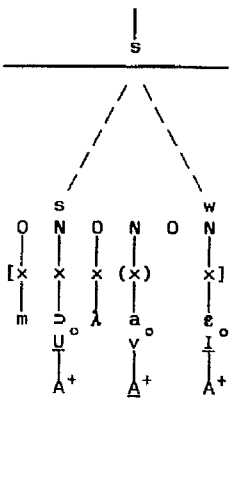
P1



[firă]

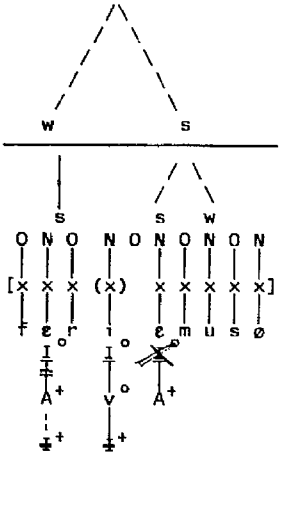


[bébă]

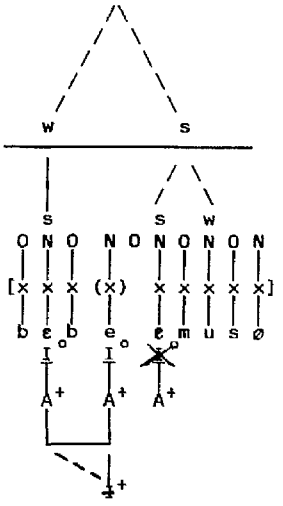


[mɔ́λĭ]

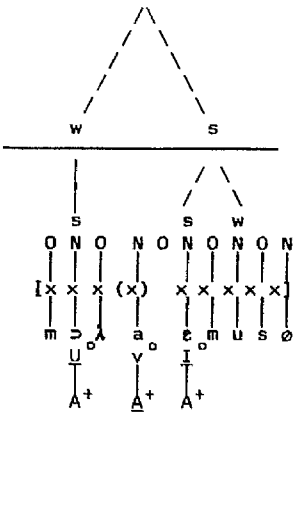
P4



[fìrámŭs]

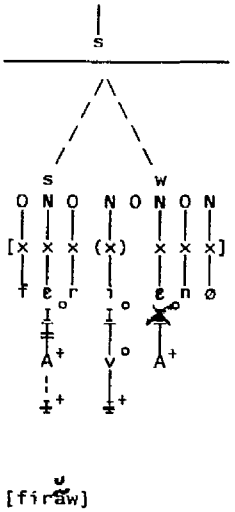


[bèbámŭs]

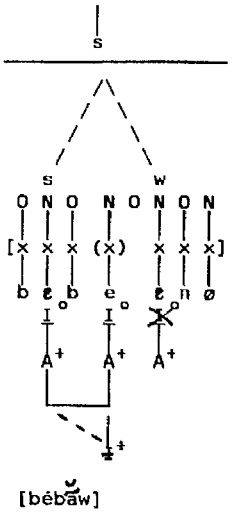


[mɔ̀λémŭs]

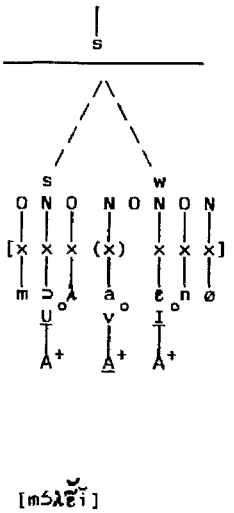
P6



[firẵw]



[bébẵw]



[mɔ́λɛ̆̃ĭ]

If we consider the realisation of the final stem nucleus with respect to the final segment in P1 above, the output strings [fírǎ] and [bébǎ] appear to be cases where vowel harmony applies when it is not obvious that it should. Conversely, [mɔ́λĭ] is a case where harmony **does not** apply although it looks as if it should. As shown in (108), whenever a sequence of two nuclei with no intervening onset point arises, the leftmost skeletal position is deleted and a floating segment (a segment which has no corresponding position on the skeletal level), the thematic vowel, results. Although the thematic vowel is floating, and hence inaudible, vowel harmony is still triggered by it. But because it does not have a skeletal point, it is not projected for metrical structure, and it follows that primary stress falls on the preceding nucleus. This type of trigger for vowel harmony entails some change in the analysis of harmony processes presented in chapter 4. In that analysis, vowel harmony is always triggered by the primarily stressed nucleus. The paradigms just analysed demonstrate that vowel harmony can also be triggered by a thematic vowel that is floating due to the loss of its skeletal position. These facts allow us to speculate on what type of nuclear segments can cause harmony. The examples show that vowel harmony can be triggered by two types of nuclei:

(109) a. Primarily stressed vowels.

b. Any floating vowel.

The cases examined show that licensed nuclear segments, i.e. segments associated to a licensed position, cannot trigger harmony. This means that only two types of unlicensed nuclear segments can exist: ((109)a) and ((109)b) above; and indeed they do both trigger vowel harmony in the Natal dialect.

5.2.1 *Successions of Nuclei With and Without an Intervening Onset Point.*

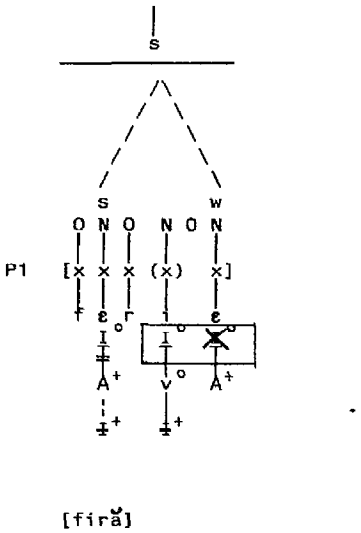
Let us now concentrate on other aspects of the derivations in (106) and (108). In the previous section we examined the processes that take place when two nuclei are adjacent in a derivation. We saw that the thematic vowel becomes floating in those circumstances. In this section we will concentrate not on the fact that the thematic vowel becomes floating when it adjacent to another nucleus with no intervening onset point, but rather, on the processes that take place at the segmental level when the same type of environment arises. Specifically, we will examine the behaviour of the various **elements** I°, U° and $A^{\dagger}$ in adjacent nuclear positions. It is therefore crucial to keep in mind that we are referring to elements (in rectangles in the derivations below) which are part of segments and not to the segments themselves.

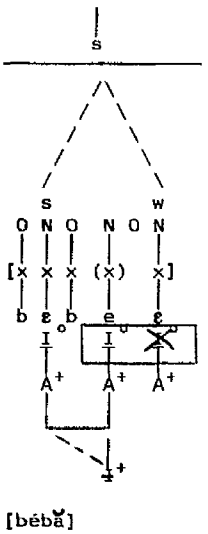
In these contexts, we notice that these elements display different patterns of behaviour when they are adjacent with no intervening onset point. Let us now consider all the possible combinations of elements in these sequences.

Let us begin by examining the I° + I° combination, which is present in the derivation of the -i- and -e- conjugations of the Subjunctive Present P1 and P4, as shown below[37]:

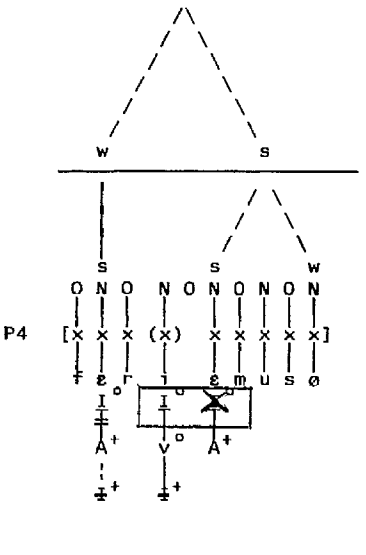
[37]In the following derivations, the deletion of an element is indicated by a large x across it.

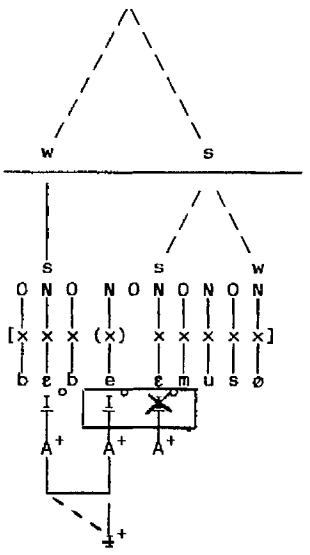
(110) Subjunctive Present





bɛbeɛ → [bébǎ]
* [bébɛ̌]
* [bébě]

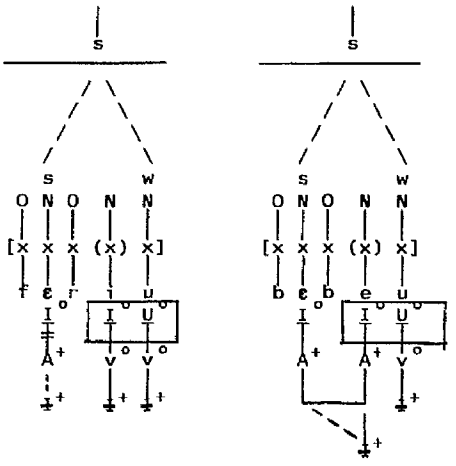




As shown above, the rightmost element deletes when two I°'s are adjacent, as confirmed by their realisation in brackets. However, it is not clear why this should be so.

Let us now consider the I° + U° sequence, which appears in the -i- and -e- conjugations of the Indicative Present P1.

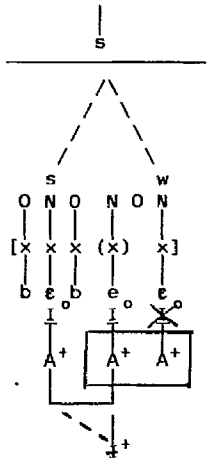
(111) Indicative Present P1.



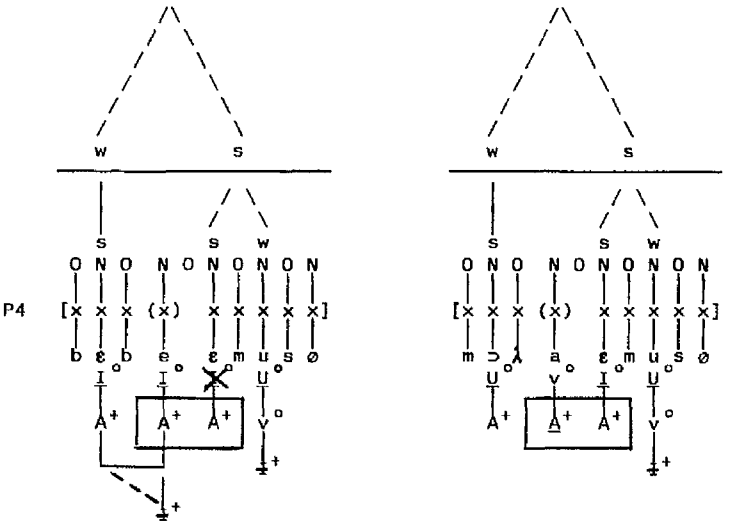
[fírŭ] [bébŭ]

As shown in (111), the adjacency of I° and U° triggers no process except for the application of OCP which leaves the thematic vowel floating, and hence inaudible. The U° element remains unaffected when it is adjacent to I°. The same happens to a sequence of two A^{+}, which occurs in the -e- conjugation of the Subjunctive Present P1, and in the -e- and the -a- conjugations of P4 in the same tense, as shown below:

(112) P1



[bébă]



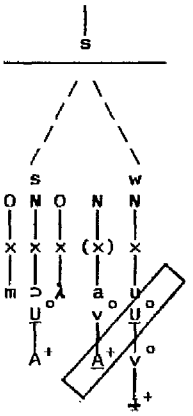
The examples above simultaneously illustrate two combination types, i.e. the adjacency of two A^+ elements, which is of primary interest here and arises in all three examples, and, coincidently, the adjacency of two I° elements, i.e. in the case of the -e- conjugation (P1 and P4). These examples illustrate, once again, that the rightmost I° element deletes when it is adjacent to another I° element. They also illustrate that the adjacency of two A^+'s triggers no deletion of elements.

So far in all cases examined, both elements in the sequence belong to the same line of representation[38]. We saw that, except for the sequence I° + I°, all elements present in the representation of the rightmost nucleus remain unchanged throughout the derivation. When two I°'s are adjacent, on the other hand, the rightmost I° is deleted.

Let us now consider the final combination, namely the combination where the elements involved belong to different lines of representation.

[38]There are no U° + U° sequences.

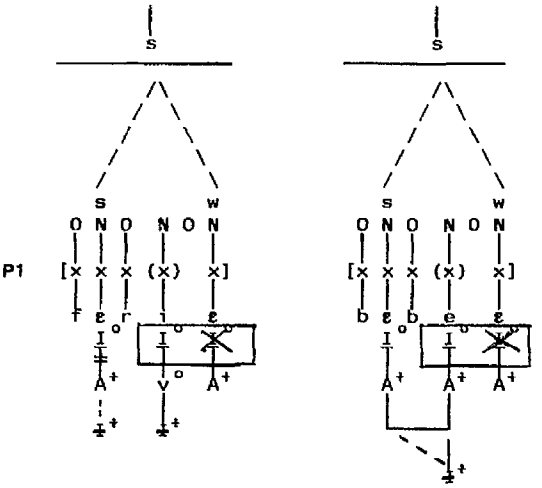
(113)



[mɔ́ʌ̀ū]

The example in (113) shows that when two elements belonging to different lines of representation are adjacent with no intervening onset point between them, the only process that takes place is the deletion of the first skeletal position, resulting in a floating TV. This allow us to conclude that the I° + I° sequence is the only case where a process occurs at the segmental level, i.e. the rightmost element is deleted. Let us go back to the I° + I° sequence (repeated below).

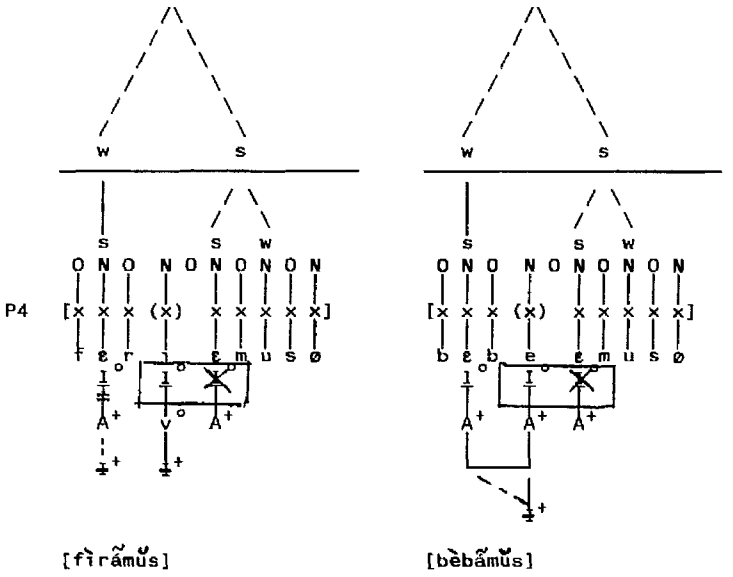
(114) Subjunctive Present



[fírə̆]

[bébə̆]

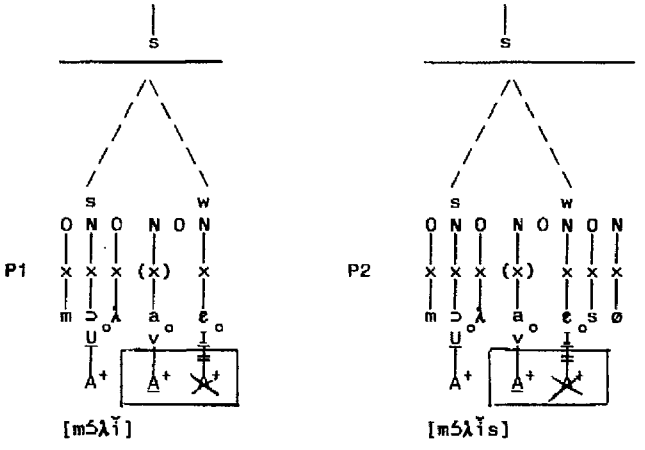
bɛbeɛ → [bébə̆]
* [bébɛ̆]
* [bébɛ̆]



According to the analysis proposed, when two I° elements are adjacent with no intervening onset point the leftmost nucleus point is lost and the thematic vowel becomes floating. Vowel harmony takes place as usual, but the rightmost I° element, for some reason, is deleted. The remaining A^{+} element is then realised as the head within that segment (Cf. their realisation).

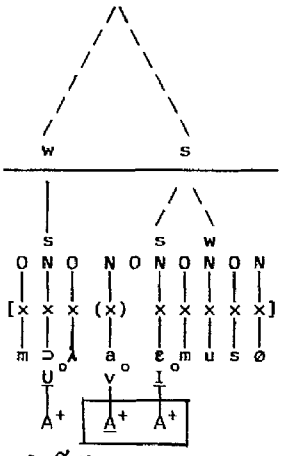
In the analysis of sequences of elements, it is important to notice that word-final non-head A^{+}'s are delinked, as shown by the derivation of the -a- conjugation of the Subjunctive Present P1 and P2 ((115)a) in contrast with that of P4 in ((115)b).

(115) a. -a- Conjugation Subjunctive Present P1 and P2.



b. -a- Conjugation Subjunctive Present P4.

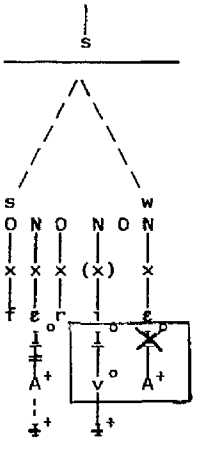
P4



[mɔ̀λ̃ẽ̆mŭs]

The deletion of post-tonic non-head A^+ is typical of most dialects of BP. In P1, the post-tonic non-head A^+ delinks. As shown in the structure of P2 in ((115)a), a post-tonic non-head A^+ will delink even if it is followed by an empty nucleus. Unlike P1 and P2, A^+ will not delink when it is primarily stressed and is followed by another full nucleus, as shown in P4. This is true for all BP dialects. The only situation where an unstressed non-head A^+ is <u>not</u> allowed to delink post-tonically occurs when the nucleus to which it belongs is part of a segment whose head has itself undergone deletion due to the application of some other process. This means that A^+ will be the only element left in the segment. This is shown below:

(116)

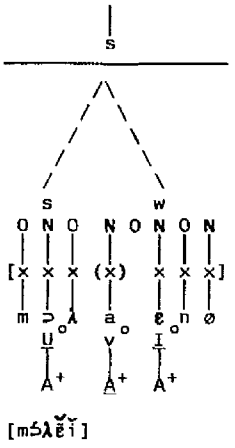


[firă] '-i- conjugation Subjunctive Present P1'

In the example above, I°, which is the head of the segment in the word final unstressed position, undergoes deletion because it is adjacent to another I°. This situation precludes the delinking of the non-head A^t because it is the only segment left to be realised in the segment in word final position.

The next example illustrates one more case where a post-tonic non-head A^t will not delink. In the following example, the nucleus where the unstressed non-head A^t belongs interacts with its neighbouring onset segment and the final empty nucleus in the formation of a diphthong (the final empty nucleus ultimately receives the I° element as a head)[39].

(117) P6



[mɔ̃λɛ̃ĩ]

The example above involves a post-tonic non-head A^t which is involved in the formation of a diphthong. It shows that in this case the delinking of A^t is blocked.

Before we can discuss the derivation of P5 in the Subjunctive Present and the Imperative, there is still one point concerning sequences of nuclei which needs some discussion. According to the analysis proposed so far, we should never find any sequences of nuclei in BP. However, these cases abound, as the realisation of the Indicative Imperfect (all persons) and P1 and P3 of the Indicative Past illustrates.

[39]The derivation of the diphthong is not discussed in this thesis.

(118) a. The Indicative Imperfect of "ferir" 'to hurt' and "beber" 'to drink'.

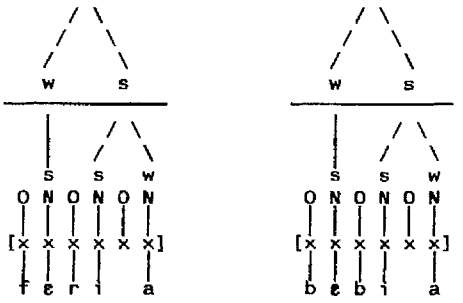
P1	[fìriă]	[bìbiă]
P2	[fìriăs]	[fìríăs]
P3	[fìriă]	[bìbíă]
P4	[fìriămŭs]	[bìbíămŭs]
P5	[fìríĕĭs]	[bìbíĕĭs]
P6	[fìríă̆ɯ]	[bìbíă̆ɯ]

b. The Indicative Past of "ferir" 'to hurt' and "beber" 'to drink'.

P1	[fìrii̯]	[bìbíi̯]
P3	[fìríŭ]	[bèbéŭ]

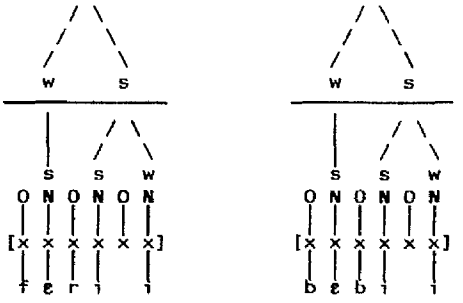
These facts require some explanation. For these structures, we propose that their lexical representation includes two nuclei which are, this time, not adjacent. They are separated from each other by an empty onset, as shown below:

(119) a. The Indicative Imperfect[40].



P1

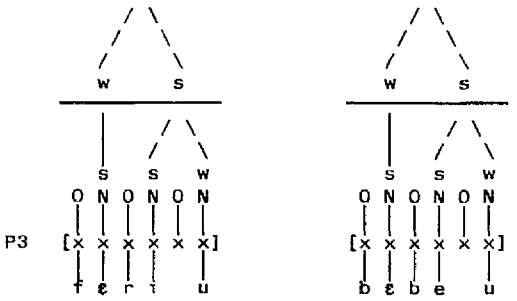
b. The Indicative Past[41].



P1

[40]The remaining persons will be found in the appendix.

[41]In the following derivations, we adopt the traditional view that the -e- conjugation takes -i- as its thematic vowel in the derivation of P1.



The structure of the Imperfect Indicative and of P1 and P3 in the Indicative Past display, as we propose, two nuclei which are separated from each other by an intervening onset point. We argue that the fact that the two nuclei are not adjacent is responsible for the non-interaction between them.

However, we need to have an independent reason to assume the existence of the onset position between the two nuclei. Evidence for this comes from the analysis of the -a- conjugation in the Imperfect Indicative, where the segment linked to the intervening onset position is audible.

(120) The -a- Conjugation Indicative Imperfect, verb "molhar" 'to shower'.

P1	[mɔ̀ʎávɐ̃]	P2	[mɔ̀ʎávɐ̃s]
P3	[mɔ̀ʎávɐ̃]	P4	[mɔ̀ʎávɐ̃mʊ̃s]
P5	[mɔ̀ʎávɐ̃ĭs]	P6	[mɔ̀ʎávɐ̃ɯ̃]

For this conjugation, we propose that the tense marker is -va-. Although all three conjugations derive from the Latin form of the person marker which had an onset, -b-, this consonant was lost in the -i- and -e- conjugations (118) and it became -v- in the -a- conjugation. The $A^{\dagger}$ element of the -e- thematic vowel was probably lost in its transition from Latin to Portuguese, since it is pronounced [i] (Cf. (118)). We will consider the loss of the $A^{\dagger}$ in the -e- conjugation shortly.

It is quite obvious that this onset segment has been maintained in the -a- conjugation simply because its absence would leave a succession of two identical nuclei, as in -mɔ̀ʎ+a+a-.

Traditionally, normative grammars assume that the -i- and -e- conjugations have merged in this tense. Here we will adopt the point of view that, in present-day BP, these conjugations do indeed share the -i- thematic vowel in a few derivations such as the Indicative Imperfect, the Past Participle, and P1 in the Indicative Past. To reinforce the view that the $A^{\dagger}$ element of the -e- thematic vowel is not lost because it is adjacent to another $A^{\dagger}$ element (in the thematic vowel), we can examine the Past Participle of the same conjugation where no such sequence exists and yet where the thematic vowel has indeed lost its $A^{\dagger}$ element, as shown below:

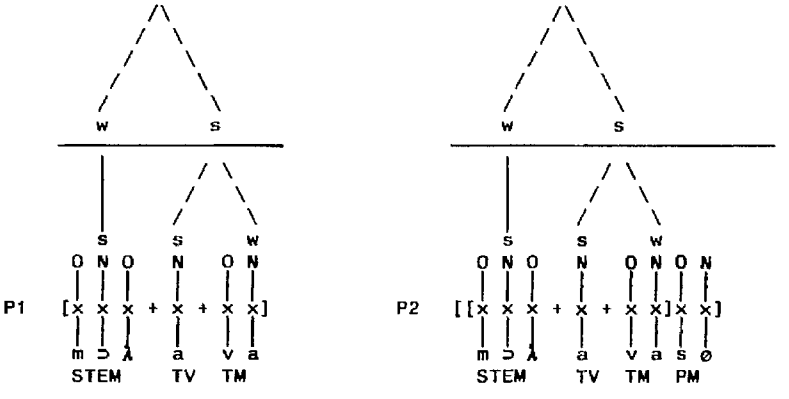
(121) Past Participle Forms.

[fìrídŭ] [bìbídŭ] (Cf. [mɔ̀ládŭ])

The fact that an onset segment intervenes between the thematic vowel and the final nucleus in this tense, and the fact that words like [fìgĕrédŭ] 'Figueiredo - surname' exist, prove that the -i- and the -e- conjugations have indeed merged in the Indicative Imperfect.

We propose the following derivation for the -a- conjugation Indicative Imperfect:

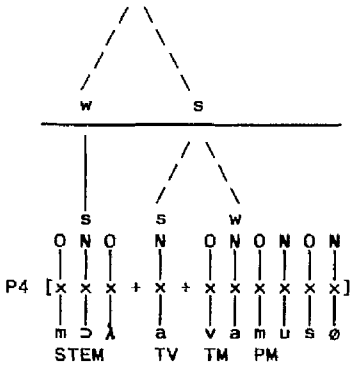
(122) -a- Conjugation Indicative Imperfect.



In the structures above, the stem is followed by the thematic vowel -a- and the tense marker -va-. Person markers, if applicable, will be added in an analytic fashion. Stress is assigned to the innermost domain. When the derivation reaches

the larger domain, stress remains as it is because the presence of nuclei to the right of an internal domain will trigger no change in its stress (Cf. chapter 3). This stage of the derivation is shown below:

(123) -a- Conjugation Indicative Imperfect.



As illustrated in (123), the person marker nucleus becomes part of the larger domain in a later stage of the derivation. However, the No-BackTracking Principle, as proposed in Chapter 3, prevents the pattern of stress assigned to the tense marker and all strings to the left of it in previous stages of the derivation from being disrupted, as it normally would. If no proviso were made, feet would be constructed from the person marker leftwards on the string. The No-Backtracking Principle is repeated below.

(124) The No-Backtracking Principle.

The phonological string is scanned from right to left. Throughout the derivation a scanning position must not vary from this direction.

Since foot construction has applied to those domains in earlier stages of the derivation, the person marker will not be accessible to the continued application of the metrical structure settings. It therefore remains unstressed.

As for the Indicative Past, the -a- conjugation is not relevant in this respect because it does not display any sequence of identical nuclei. Because its structure is to some extent similar to the derivation of P5, it will be discussed along with the P5 marker below.

The Imperfect and P1 and P3 of the Indicative Past are not the only verb forms which exhibit a succession of audible nuclei in BP. The structure of P5 in most tenses, although masked by the effects of other phenomena that operate in each tense (which will be examined below) illustrates the existence of adjacent nuclei in BP, as shown below:

(125) Affirmative Imperative:

[fiɾiĭ] [bèbéĭ] [mɔ̀ʎáĭ]

Just as we proposed for the preceding cases, the two nuclei in question are in reality separated from each other by an intervening onset point, which has no phonological content[42]. This means that neither of the nuclei will undergo any sort of deletion process or become floating as in the cases examined so far.

We have independent evidence from the realisation of the Personal Infinitive to support the presence of this onset position. In the Personal Infinitive and its homonymous Subjunctive Future, the P5 marker is realised as -dis-, as shown below:

(126) The Personal Infinitive[43].

[fìr	+	i	+	r	+	dis]	[bèb	+ e +	r + dis]
STEM		TV		TM		PM	STEM	TV	TM PM
[mɔ̀ʎ	+	a	+	r	+	dis]			
STEM		TV		TM		PM			

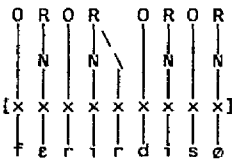
We propose that these forms derive from the following structures:

---

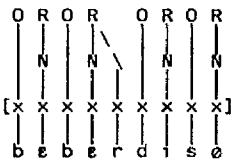
[42]The actual derivation of the Affirmative Imperative P5 will be discussed below.

[43]The actual realisation of the tense marker (TM) -r- as a velar fricative is not relevant for the discussion.

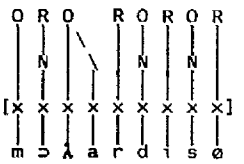
(127) The Personal Infinitive.



[firihdĭs]
"ferirdes"
(for you) to hurt'



[bèbéhdĭs]
"beberdes"
'(for you) to drink'

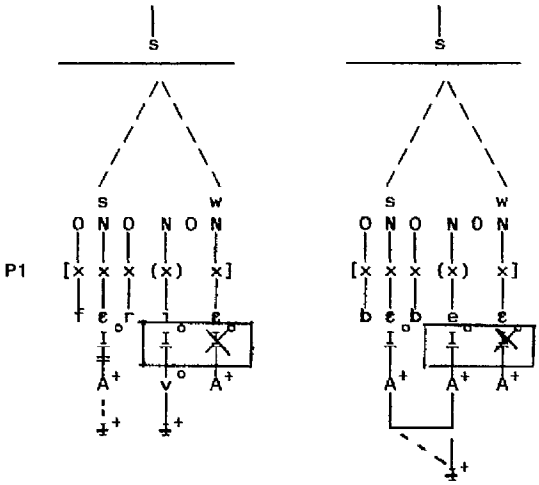


[mɔ̀ʌ́áhdĭs]
"molhardes"
(for you) to shower'

Let us recapitulate the point of this discussion. We have claimed that the Subjunctive Present and most persons of the Imperative are characterised by a sequence of two nuclei with no intervening onset position. We claimed that the adjacency of those two nuclear positions is responsible for the series of processes that take place in those structures and causes the thematic vowel in each case to become floating. We showed that the presence of an empty onset with a point is enough to maintain the autonomy of the two otherwise adjacent nuclei. Specifically, we showed that in the -i- and the -e- conjugations of the Indicative Imperfect and in P5 of most tenses, the presence of an intervening onset point prevents the two nuclei from interacting, as they do in the Subjunctive Present and in the Imperatives, where no point is present[44]. In the analysis of the Subjunctive Present and in the Imperatives, as well as in P1 Indicative Present, we noticed that when two I° elements are part of the representation of strictly adjacent nuclei, the rightmost one deletes ((110), repeated below).

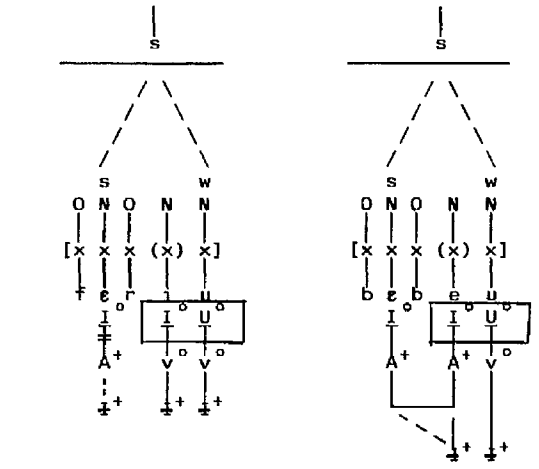
[44]Except for P5 in the Affirmative Imperative.

(128) Subjunctive Present

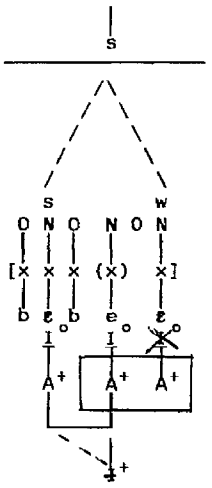


Other clusters, like I° + U° and A^{+} + A^{+}, are maintained in the same circumstances.

(129) a. Indicative Present P1.



b. Subjunctive Present P1



[bébă̆]

In this section, we examined cases where two nuclei are separated from each other by an empty onset. We observed that the presence of the empty onset is enough to prevent any interaction between the two otherwise adjacent nuclei. We gave evidence for the presence of the empty onset position with the analysis of the Affirmative Imperative P5 and the Personal Infinitive and of the -a- conjugation in the Indicative Imperfect. In these derivations, the segment which is linked to the onset position in each case is actually audible.

5.2.2 *The Derivation of P5.*

Our discussion of the adjacency of nuclei brings other types of phenomena into consideration. Specifically, the derivation of P5 in many tenses is accompanied by some less obvious phenomena.

Let us consider the representation of P5 in detail. We will first compare the realisation of P5 in the Subjunctive Present, in the Affirmative and the Negative Imperatives, which is given below, and then extend this analysis to P5 in the remaining tenses:

(130) a. Affirmative Imperative:

[fìriĭ̆] [bèbéĭ̆] [mɔ̀λáĭ̆]

b. Negative Imperative:

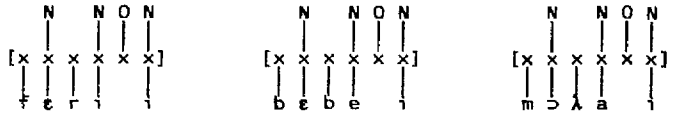
[fìráĭs] [bèbáĭs] [mɛ̀lɛ́ĭs]

c. Subjunctive Present:

[fìráĭs] [bèbáĭs] [mɛ̀lɛ́ĭs]

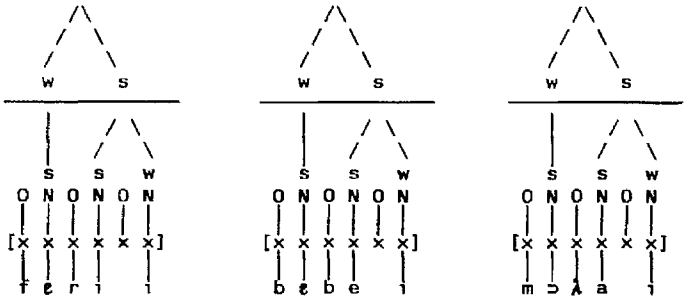
From the fact that in these cases the final stem nucleus is not primarily stressed and is harmonised according to the vowel harmony analysis presented in chapter 4, we can conclude that the thematic vowel is present in the lexical representation of all three tenses. In fact, the thematic vowel is clearly audible and receives primary stress in the Affirmative Imperative ((130)a). It is also clear that it has become floating via the application of the same type of processes discussed in section 5.2.1 in the Subjunctive Present and the Negative Imperative. We already know that two nuclei will not interact if a skeletal onset point intervenes between them. This observation allows us to conclude that the lexical representation of the Affirmative Imperative is as follows:

(131) The Affirmative Imperative.



The metrical analysis of this tense may then be undertaken, as follows:

(132) The Affirmative Imperative.

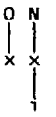


As shown in (132), the thematic vowel and the person marker are separated by an empty onset. For this reason, the two nuclei do not interact. Stress is then assigned as proposed throughout this

thesis. This analysis predicts the attested pattern of stress, as shown in ((130)a).

We propose that P5 in this tense is lexically represented as follows (Cf. its representation in (134)):

(133) P5 in the Affirmative Imperative:

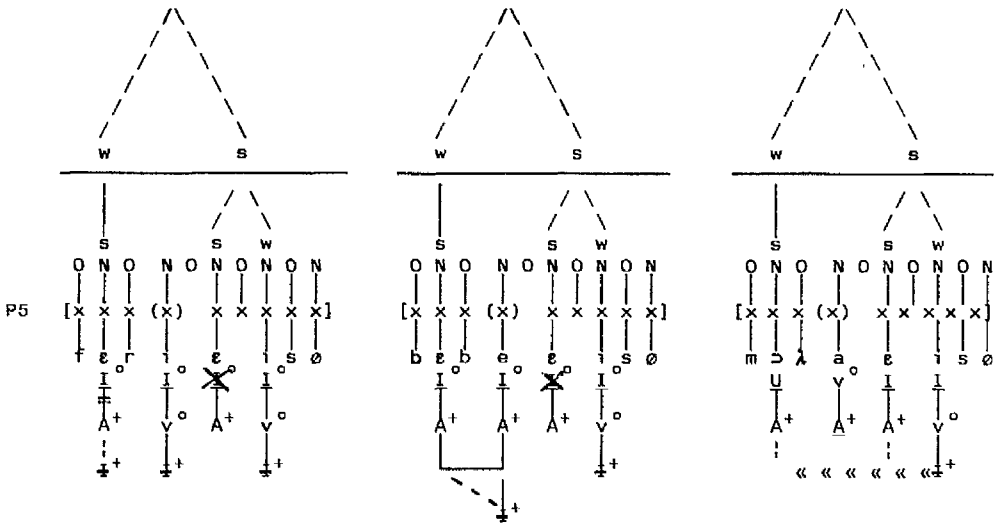


As for the other two derivations in (130), we know that they have -ɛ- as their tense marker, given the realisation of the final stem nucleus in the -i- conjugation in the Subjunctive Present, [fi̱ráĭs]. In this conjugation, the I° element is deleted because it is adjacent to another I° element in the thematic vowel, as seen in other cases. As for the person marker, we propose that in addition to the string proposed for the Affirmative Imperative proposed in (133), it also has an -s-, as illustrated below:

(134) P5 in the Subjunctive Present and Negative Imperative:

We propose that the Subjunctive Present and the Negative Imperative, which surface in exactly the same way, are derived from the following structure:

(135) a. The Subjunctive Present and Negative Imperative.



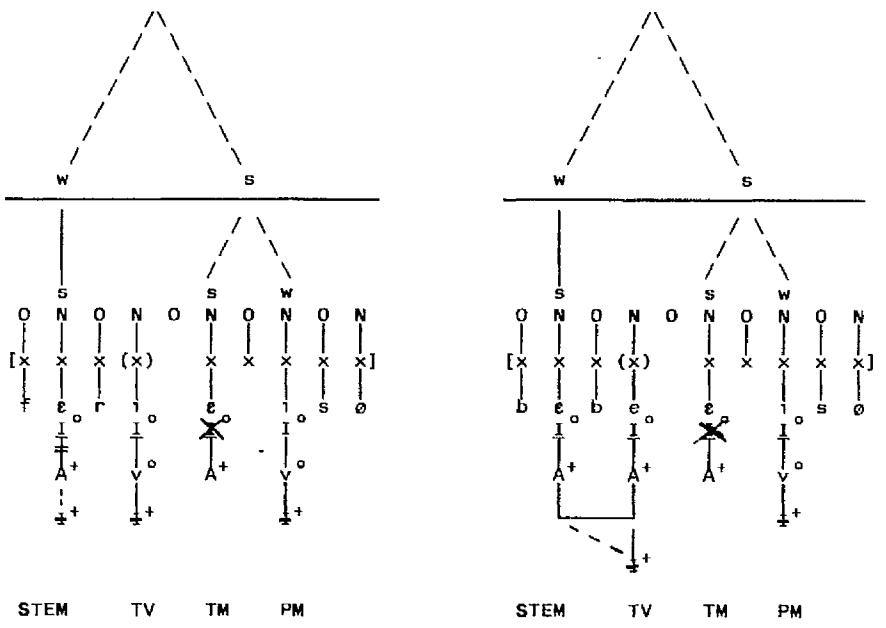
[firáĭs]
"firais"
'(you) to hurt'

[bèbáĭs]
"bebais"
'(you) to drink'

[mòʎéĭs]
"molheis"
'(you) to shower'

Let us now consider the derivation of these structures step by step. At a glance, we can see that they share at least two steps in the derivation. The first stage which is common to both structures is the fact that the thematic vowel becomes floating due to its adjacency to the tense marker nucleus with no intervening onset point. The second is that in all cases the person marker starts with an empty onset with a point. This means that the tense marker nucleus, -ɛ-, will not be adjacent to the following (person marker) nucleus as it is to the preceding thematic vowel. Since the tense marker nucleus and the person marker nucleus are separated from each other by an empty onset with a point, they are not expected to interact. Let us now consider each conjugation individually, starting with the -i- and -e- conjugations, whose structures are repeated below:

(136)

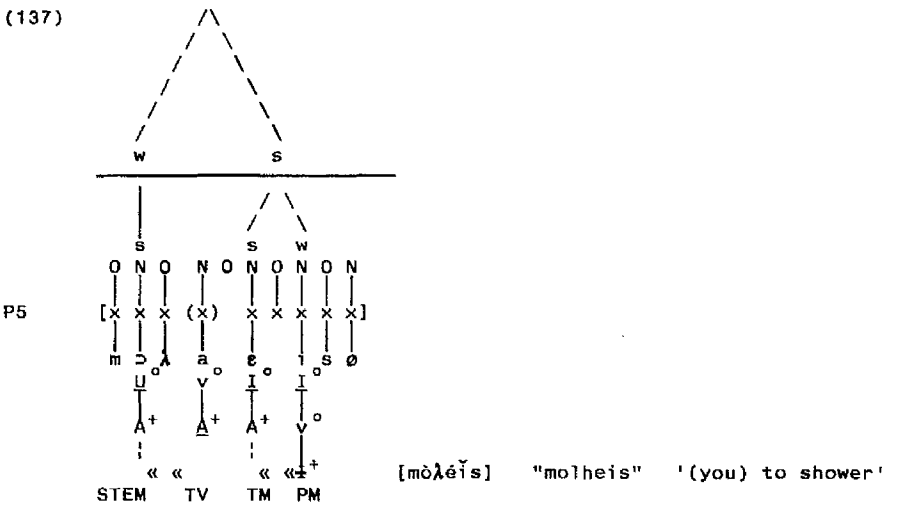


As shown above, in the -i- and -e- conjugations, the thematic vowel skeletal point is delinked. As we have seen in the preceding cases of adjacency of two nuclei with no intervening nucleus, the thematic vowel becomes floating. Vowel harmony applies as proposed in chapter 4. The unlicensed A^+ element in the stem nucleus in the -i- conjugation is deleted. The ɪ^+ element becomes automatically available in that nucleus.

These conjugations are also characterised by the adjacency of I° elements in the thematic vowel and in the tense marker. As we saw in the discussion on (116), whenever two I°'s become adjacent the rightmost I° element delinks. In both cases, the attested outcome is exactly as illustrated in ((130)b) and ((130)c), [fìráĭs] and [bèbáĭs].

Let us now check whether the same thing happens in the -a- conjugation.

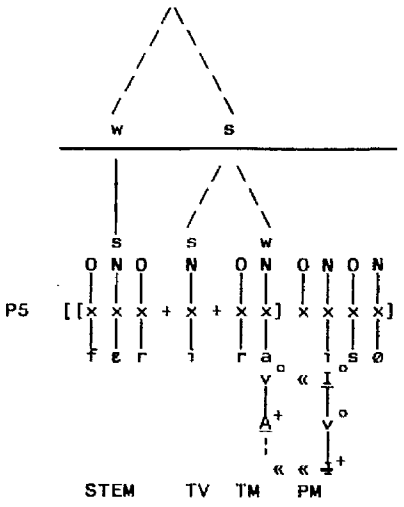
(137)

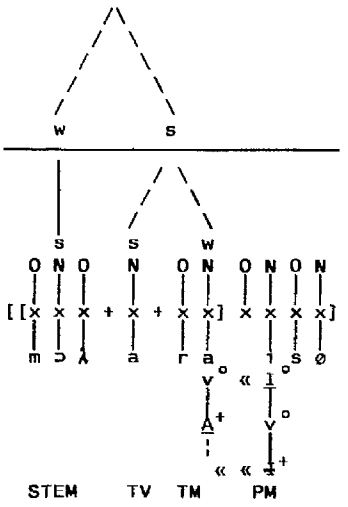


In this derivation, the thematic vowel becomes floating because it is adjacent to the tense marker nucleus, as mentioned above. According to our argument so far, no further processes are expected, which means that the realisation of this string should be [mɔ̀λɛ́ĭs]. However, as illustrated in ((130)b) and in ((130)c), the tense marker nucleus is realised as a tense mid vowel, which suggests that spreading of ɪ^{+} is taking place. As far as can be seen from the structure in (137), the only possible source for the spreading is the tense marker nucleus -i-, across the empty onset, which then changes the tense marker -ɛ- into [e], as illustrated in the attested form, [mòλéĭs]. Although the actual reason why spreading of ɪ^{+} occurs in these circumstances remains unclear, this discussion predicts that the outcome should be one where the stem nucleus is lax, as in [mɔ̀λéĭs]. The attested realisations in ((130)b) and in ((130)c) show that the spreading of ɪ^{+} does not stop at the floating thematic vowel, but continues leftwards. Spreading takes place across the floating thematic vowel and makes the stem nucleus tense as well. The surface form [mòλéĭs] results.

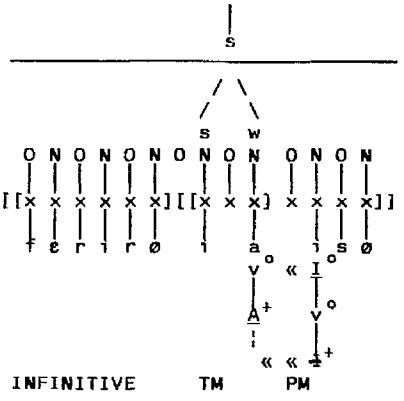
The same type of derivation illustrated above also takes place in the derivation of P5 in most tenses, as illustrated below:

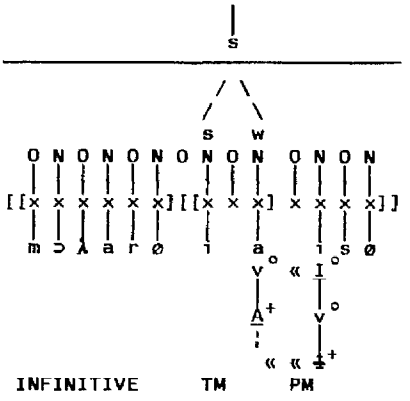
(138) a. Pluperfect.





b. The Conditional.

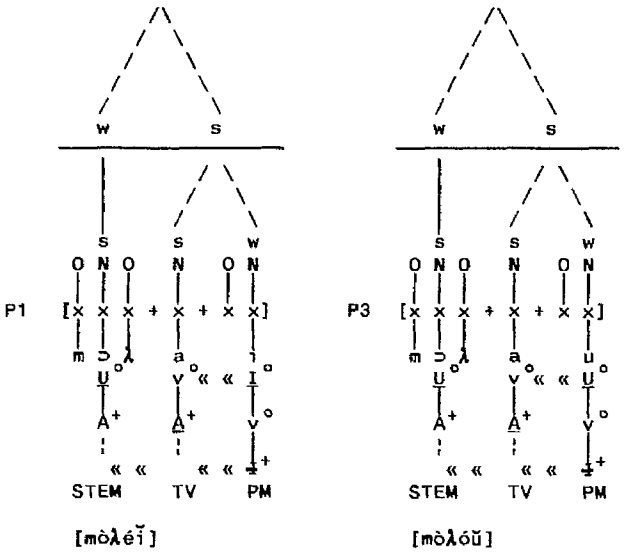




As illustrated by the examples above compared with those in (136), not only ɪ⁺ but also I° spreads onto the tense marker nucleus. The spreading of I° and ɪ⁺ occurs whenever the P5 marker (illustrated in (134)) follows a domain final -a-.

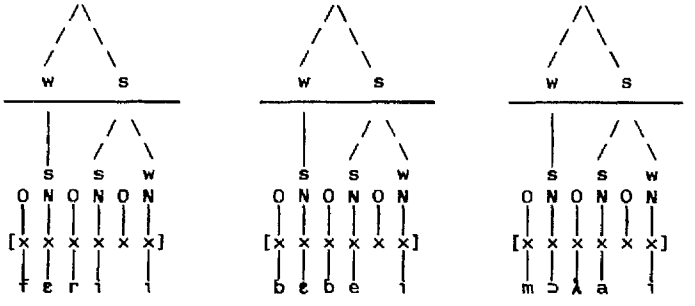
The analysis would be perfectly comprehensive if spreading of I° and ɪ⁺ would only occur onto a domain final -a-. However, spreading also occurs in some unexpected contexts, for example in the following derivations:

(139) -a- Conjugation Indicative Past.



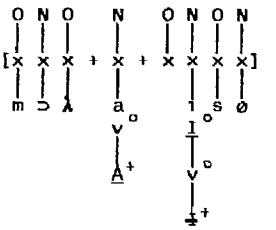
The derivation of the Indicative Present P1 and P3 of the -a- conjugation as illustrated above is problematic for two reasons. Not only do I° and ɬ$^{+}$ spread onto a non-domain final -a-, but also, the lexical representation of P1, as it stands, coincides with the lexical representation of P5 in the Affirmative Imperative (Cf. (132), repeated below).

(140) The Affirmative Imperative.



On the other hand, the derivation of P5, which displays the same type of structure as P1, except for the presence of the Indicative Present P5 marker, poses no problem for our analysis. The derivation of P5 in this tense is illustrated below:

(141) a. Indicative Present.



[mɔ̀lǎĭs]

Let us sum up our results so far. We have analysed several verb paradigms whose structures are less obvious than those seen in the preceding chapters. We have seen that the derivations of the Subjunctive Present, and of the Affirmative and the Negative Imperatives are characterised by a sequence of two adjacent nuclei with no intervening onset point. We proposed that other tenses, such as the Indicative Imperfect, P1 and P3 Indicative Past, and the P5 marker, display a sequence of two nuclei which are separated by an empty onset with a point. We have seen that these two types of sequences display considerably different behaviour. The adjacency of two nuclei with no intervening onset point entails the deletion of the leftmost skeletal position ((111) and similar structures), whereas two nuclei which are separated by an onset which has a point maintain their autonomy (119). We have also noted that the sequence I° + I° entails deletion of the leftmost I° (110).

Besides these aspects of the lexical representation of verb paradigms, we discovered a new type of trigger for vowel harmony, namely a floating segment. We also examined the spreading of elements from the person marker nucleus onto the preceding one. We noticed that spreading of I° and ɪ^{+} from the P5 nucleus normally occurs onto a domain final -a-, with two exceptions, i.e. the derivation of the Indicative Past P1 and P3. The diversity of facts which occur in each type of structure can undoubtedly be considered the best evidence for the structures proposed.

The analysis presented in this section can successfully account

for the long-assumed "odd facts" of the Subjunctive Present and the Imperatives. Even the most recent analyses available for these derivations included the following type of explanation for these facts:

(142) "Person marking is supplemented by metaphonic alternations in the second and third conjugations[45]... The general effect of metaphony is that the distinction between low-mid and high-mid root vowels is eliminated. The only exceptions to the alternations are verbs where the root vowel quality is invariable or where mid-vowel distinctions are neutralised by a basic phonological rule. In abstract morphophonemics it is possible to account for these alternations by metaphonic rules assimilating the first person singular root vowels to the thematic vowels, thus giving high vowels when the theme is /i/, high-mid when it is /e/, and low when it is /a/." (Parkinson, 1988).

These analyses are only a description of the circumstances where the "raising" occurs. There was no direct link between the process and the context in which it occurred. Given our analysis, where phenomena are directly motivated by their context, the "raising" phenomenon of the primarily stressed stem nucleus in these derivations in the -i- and the -e- conjugations is now to a great extent explained.

It is important to note that, for the Natal speaker, the realisation of pretonics in the Subjunctive Present, the Imperatives and P1 Indicative Present and the pretonics analysed in chapter 4 represent exactly the same phenomenon. Specifically, they all result from the application of vowel harmony processes and they represent the only possible outcome. For Southern dialect speakers, on the other hand, this is not true since pretonics are invariably complex and tense, as shown below:

(143) a. Imperfect

[ʒɛ̲lávă]	* [ʒɛ̲̀lávă]	'(it) would get cold'

b. Indicative Present

[ʒé̲lŭ]	* [ʒélŭ]	'(I) get cold'

According to the analysis presented, there is no immediate source for the tenseness of the pretonic in ((143)a). Also, as far as the Indicative Present is concerned, we cannot understand why the

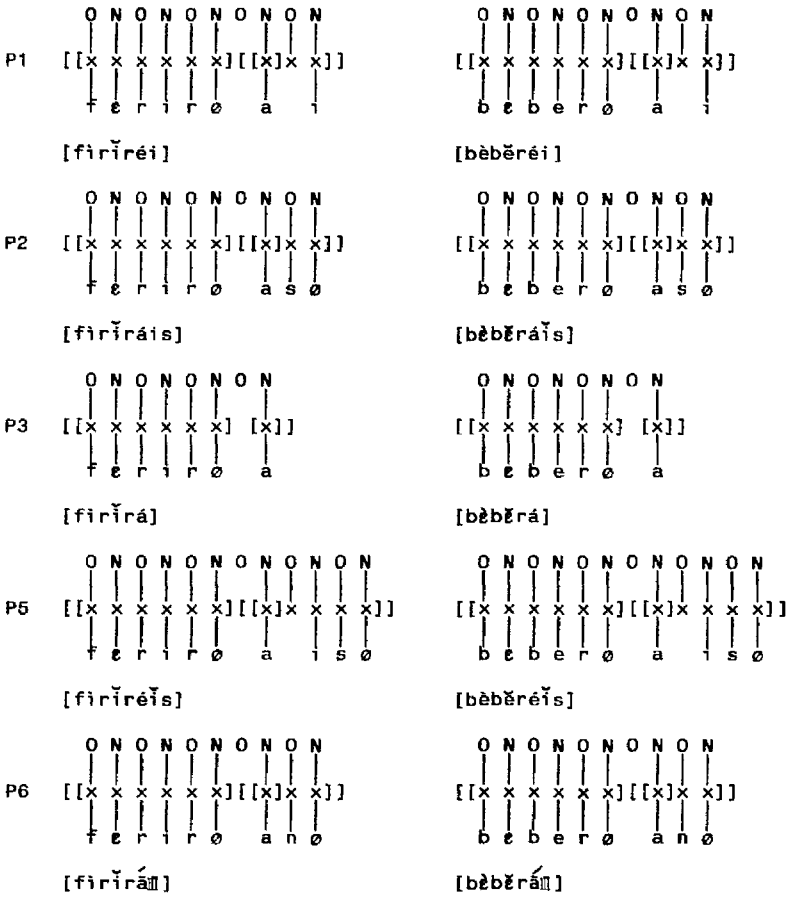
[45]Second and third conjugations in this work refers to our -e- and -i- conjugations.

primary stressed nucleus is obligatorily lax. The only way we can explain these facts is if we assume that Southern dialects display only the floating type of vowel harmony trigger[46]. ((143)b) is then accounted for.

5.2.3 *Cases Which Cannot Be Accounted for by This Analysis.*

Throughout the analysis of verb paradigms, the derivation of the Indicative Future P4 was left untouched. That is because its derivation does not follow the pattern proposed for this tense, as can be compared below:

(144) Indicative Future



From the structures proposed for the Indicative Future forms and

[46]As far as the examples analysed here are concerned.

their realisation proposed in (144), we should conclude that the same type of structure would be present in P4, namely STEM+TV+TM+PM, whose the PM is -mus. However, for reasons which thusfar defy adequate explanation, in P4, an extra I° element is added to the tense marker nucleus. The paradigms are then realised as follows:

(145) Indicative Future P4.

[fiɾĭɾɛ̃mŭs] [bɛ̀bɛ̆ɾɛ̃mŭs] [mɔ̀lăɾɛ̃mŭs]

According to the analysis proposed for verb paradigms so far, there seems to be no way of deriving the results in (145). If we propose that the extra I° element is added either within the same domain or outside the domain where the tense marker nucleus -a- is with no intervening onset position, according to the analysis of the Subjunctive Present and Imperatives, the tense marker nucleus -a- should not surface. In that case, the realisations of this form should be * [fiɾĭímŭs], * [biбĭímŭs] and * [mɔ̀lăɾímŭs], which are not attested. On the other hand, if we propose that the tense marker -a- and the I° element are separated from each other by an onset point, the expected outcome should be * [fiɾĭíămŭs], * [biбĭíămŭs] and * [mɔ̀lăɾíămŭs], which are likewise unattested. This means that we are unable to derive the Indicative Future P4 following the analysis proposed in this thesis.

This concludes our discussion on the less straightforward derivations in verb paradigms of BP.

6.1 *Introduction.*

This chapter, which opens the discussion on the assignment of stress to nouns and adjectives and related phenomena (chapters 6 through to 9), deals with nouns and adjectives which display two of the three patterns of stress which are possible in these grammatical categories, i.e. the penultimate and the final patterns of stress. The third pattern, namely antepenultimate primary stress, will be dealt with in chapter 9. In this discussion, we will attempt to prove that these two grammatical categories are stressed with the same set of principles and parameters proposed for the assignment of stress to verbs.

We start off with a discussion of some morphological aspects of nouns and adjectives which we believe to be crucial for the analysis of stress we propose. After that, we extend this discussion to an understanding of the assignment of stress to these two grammatical categories.

6.2 *The Morphology of Nouns and Adjectives in BP.*

Since we are grouping together nouns and adjectives, we start with a description of the general characteristics of each of these grammatical categories.

Nouns and adjectives in BP are traditionally assumed to display the general morphological structure shown below:

(146) STEM + THEMATIC VOWEL (TV)

Here we will adopt this view and, because the TV generally reflects the gender of the lexical item, we will refer to it as the gender suffix (henceforth, GS).

This general morphological characteristic is shown in the following examples, where the GS is underlined:

(147) a. Nouns[47]

[ká:hŭ]	"carro"	'car'
[vê:lă]	"vela"	'candle'
[lá:sŭ]	"laço"	'bow'
[mé:ză]	"mesa"	'table'

b. Adjectives

[bùní:tŭ]	"bonito"	'beautiful (masc)'
[kàlá:dă]	"calada"	'silent (fem)'
[pìkẽ́:nŭ]	"pequeno"	'small (masc)'
[bàrá:tă]	"barata"	'cheap (fem)'

As shown in (147), nouns and adjectives in general display the structure in (146), i.e. the final nucleus reflects the gender of the lexical item. As a rule of thumb, [a] -a is used for the feminine and [u] -u for the masculine. As we will see, all exceptions to this pattern are learned on a one-to-one basis. We will discuss all possibilities below.

Even though we are treating nouns and adjectives as a single category, we have to admit that the spell-out (i.e. realisation) of the GS in these two categories have different status. While the gender in nouns is clearly an idiosyncratic property of individual lexical items, the spell-out of the gender in adjectives generally reflects the gender of the noun it complements, whether the noun is present or simply understood:

(148)	[ká:hŭ]	[bùní:tŭ]	"carro bonito"	'beautiful car'
	[vê:lă]	[bùní:tă]	"vela bonita"	'beautiful candle'
	[lá:sŭ]	[pìkẽ́:nŭ]	"laço pequeno"	'small bow'
	[mé:ză]	[pìkẽ́:nă]	"mesa pequena"	'small table'

The examples in (148) show that the GS in adjectives depends on the gender of the noun it complements. At a later stage in the discussion we will consider both the cases where the GS does not reflect the gender of the noun it complements, and the interpretation of this analysis in terms of lexical representation.

So far we have seen the basic cases of gender spell-outs in nouns and in adjectives. We have noticed that their basic morphological structure is STEM+GS. We have also seen that the general pattern

[47]The lengthening of the primary stressed nucleus, which is irrelevant to the discussion, is only discussed in chapter 9.

for the spell-out of the GS is that [a] -a is used for the feminine and [u] -u is used for the masculine, and that although it is a lexical property in nouns, in adjectives it usually reflects the gender of the noun it complements.

We now move on to some less straightforward cases. There are cases where the GS spell-out does not follow the basic pattern. Some lexical items simply have no gender suffix at all, as shown below[48,49]:

(149) a. Nouns

[sìpɔ́:]	"cipó"	'stick'	(Masc)
[màré:]	"maré"	'tide'	(Fem)
[sàbó:]	"sabor"	'flavour'	(Masc)
[fló:]	"flor"	'flower'	(Fem)
[nàtám]	"Natal"	'Christmas'	(Masc)
[mɔ̀rám]	"moral"	'morale'	(Fem)
[méjs]	"mês"	'month'	(Masc)[50]
[fɔ́js]	"foz"	'source'	(Fem)
[àrmãzẽj]	"armazem"	'warehouse'	(Masc)
[nàsãm]	"nação"	'nation'	(Fem)

b. Adjectives

[vùmgá:]	"vulgar"	'vulgar'	(Masc/Fem)
[sùtím]	"sutil"	'subtle'	(Masc/Fem)
[hùĩ]	"ruim"	'bad'	(Masc/Fem)
[kàpájs]	"capaz"	'capable'	(Masc/Fem)

As seen above, words such as these carry no mark to indicate their gender[51].

---

[48]In the next cases, since the gender does not follow the general rule of -a for the feminine and -u for the masculine, the gender information will be given in parentheses after the translation.

[49]The derivation of diphthongs from sequences of consonants followed by an empty nucleus is not discussed in this thesis.

[50]This type of diphthong is typical of Natal.

[51]We will come back to the details of the lexical representation of each of these words at a later stage in the discussion.

Other words, instead of having [a] -a or [u] -u as the spell-out of their gender, have [i] -e. These cases are illustrated below:

(150) a. Nouns

[sidá:dĭ]	"cidade"	'city'	(Fem)
[àláhmĭ]	"alarme"	'alarm'	(Masc)
[pá:sĭ]	"passe"	'pass'	(Masc)
[pã̂:nĭ]	"pane"	'fault'	(Fem)
[bú:lĭ]	"bule"	'coffee-pot'	(Masc)
[pé:lĭ]	"pele"	'skin'	(Fem)

b. Adjectives

[vàlẽ̂tĭ]	"valente"	'brave'	(Masc/Fem)
[sɔ̀lẽ̂:nĭ]	"solene"	'solemn'	(Masc/Fem)
[prèkɔ́:sĭ]	"precoce"	'early developer'	(Masc/Fem)
[mɔ́:lĭ]	"mole"	'soft'	(Masc/Fem)

The final group of less straightforward cases includes a few words whose gender suffix spell-out is either the reverse of the general pattern, or has only one form for both the feminine and the masculine. This is the case in only a small number of words, as illustrated below:

(151) a. Nouns

[mɐ̃̂ɯ]	"mão"	'hand'	(Fem)
[gwáhdă]	"guarda"	'warden'	(Masc/Fem)
[èrĕmí:tă]	"eremita"	'hermit'	(Masc/Fem)
[àhtístă]	"artista"	'artist'	(Masc/Fem)

b. Adjectives

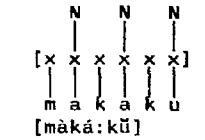
[ìdĭɔ́:tă]	"idiota"	'idiot'	(Masc/Fem)
[sìprĭɔ́:tă]	"cipriota"	'native of Cyprus'	(Masc/Fem)

As shown in (151), there are feminine words whose GS position is filled with -u, as illustrated by [mɐ̃̂ɯ]. In other cases (as shown in the other examples), the spell-out of the GS is -a even though the word can refer to the masculine.

Before we continue, it is crucial to highlight a notational distinction. We need to distinguish cases where, by alternating [a] -a and [u] -u we can generate the gender in a noun or adjective, from cases where a single ending for the GS position applies indiscriminately to either masculine or feminine. We propose to make this distinction by specifying (Masc/Fem) for cases where the alternation -a/-u produces the two genders, and Free Choice (henceforth, FC) for cases where one single ending

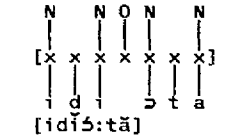
applies to both the masculine and the feminine. This distinction is illustrated by the following examples:

(152) a.



[màká:kŭ]
"macaco"
'monkey'
(Masc/Fem)

b.



[idĭɔ́:tă]
"idiota"
'idiot'
(FC)

The information concerning the gender of each of these forms is defined in the following way. Whilst the derivation in ((152)a) can yield two possible outputs, [màkákŭ] for the masculine and [màkákă] for the feminine (we will discuss these derivations below), the derivation in ((152)b) yields only [ìdĭɔ́:tă] for both the feminine and the masculine. This shows how these two cases will be distinguished in this thesis. There is still a further possibility, in which only one gender exists. These cases will be indicated by the GS content (and the absence of lexical information concerning the gender). Following this convention, the words in ((149)b), ((150)b) and two cases of (151) now need to be updated to the following:

(153) Adjectives

[vùm̃gá:]	"vulgar"	'vulgar'	(FC)
[sùtíl̃]	"sutil"	'subtle'	(FC)
[hùĭ̃]	"ruim"	'bad'	(FC)
[kàpájs]	"capaz"	'capable'	(FC)
[vàlẽ̆tĭ]	"valente"	'brave'	(FC)
[sɔ̀lẽ̆:nĭ]	"solene"	'solemn'	(FC)
[prɛ̀kɔ́:sĭ]	"precoce"	'early developer'	(FC)
[mɔ́:lĭ]	"mole"	'soft'	(FC)
[gwáhdă]	"guarda"	'warden'	(FC)
[ìdĭɔ́:tă]	"idiota"	'idiot'	(FC)

The cases where the derivation yields only one gender is illustrated below:

(154)	[mé:ză]	"mesa"	'table'
	[ká:hŭ]	"carro"	'car'
	[vé:lă]	"vela"	'candle'
	[lá:sŭ]	"laço"	'bow'

The absence of any extra lexical information means that it is not possible to make the masculine/feminine alternation in words such as these.

Let us now continue with the discussion on the gender suffix. So far we have seen cases where morphology consists maximally of a stem followed by a gender suffix. We have seen that in some cases the gender suffix is not present. However, we need to give evidence for the absence of the GS in cases where it does not surface. Specifically, we need to show that the primary stressed nucleus in [fló:] and [sàbó:], for example, is not a gender suffix. To support this analysis, we will base our argument on the following assumption. We propose that a gender suffix is present at the end of a domain only. Whatever precedes or follows the stem, only one GS will be present, i.e. at the end of the domain. We propose, furthermore, that whatever is part of a stem cannot be lost in the course of a derivation. This being the case, let us now consider words with a more complex morphological structure.

Consider the following words:

(155)	[mé:ză]	[mèzĩ:ñă]	[mèzõ:nă]
	"mesa"	"mesinha"	"mesona"
	'table'	'little table'	'big table'
	[lá:sŭ]	[làsĩ:ñŭ]	[làsá:dă]
	"laço"	"lacinho"	"laçada"
	'bow'	'little bow'	'knot'
	[àlármĭ]	[àlărmá:dŭ]	[àlărmãtĭ]
	(Masc)	(Masc/Fem)	(FC)
	"alarme"	"alarmado"	"alarmante"
	'alarm'	'alarmed'	'alarming'
	[fló:]	[flòrí:dŭ]	[flòrístă]
	(Fem)	(Masc/Fem)	(FC)
	"flor"	"florido"	"florista"
	'flower'	'flowered'	'florist'

As shown above, a stem maintains its integrity throughout a derivation. So, in a word like [mé:ză], the stem -mez- remains intact when affixes are added to form other words, while the

gender suffix is omitted. In cognitive terms, we are proposing that when the speaker of BP acquires vocabulary, he/she learns that although lexically the GS position is filled with some segment or, in the cases where there is no GS, it is absent, this position must be dropped in order to derive other words. The speaker learns not only how to parse words morphologically but he/she also learns that there are very productive word formation "rules" which can be used to derive other words. We will come back to the interpretation of this analysis in cognitive terms as we discuss the lexical representation of nouns and adjectives (below).

This means that in the derivation of 'little table' (155), for example, we will have a stem to which -iñ- is added before a GS position can be added. The newly derived word has a derivational stem (=a string consisting of a stem plus affixes) followed by a gender suffix, domain finally. The same process applies to 'big table'. The initial stem, -mez-, is concatenated with -ɔna- where -a- is the spell-out of the GS. A GS will be found at the end of the domain, unless the final affix does not end in one. This process, we propose, goes on for as long as the concatenation of affixes is morphologically possible. The combination of the various affixes is, of course, determined by the morphology.

However, the derivations in (155) need further discussion. In certain cases, for instance in the case of "laço" and "alarme", the derivation ends up with a GS spell-out which is not appropriate to the stem's initial requirement. We propose that this is due to the fact that, in BP, the final suffix determines what segment, if any, can occupy the gender suffix position, even though the content might in some way be semantically linked to the initial stem. What this means is that morphologically the final suffix is responsible for the GS that follows it. This is shown below:

(156)	[bànɐ̃́:nă]	[bànănáːdă]	[bànănéjră]
	"banana"	"bananada"	"bananeira"
	'banana'	'banana juice'	'banana tree'

[àbăká:tĭ]	[àbàkătá:dă]	[àbàkătéjrŭ]
(Masc)		
"abacate"	"abacatada"	"abacateiro"
'avocado'	'avocado juice'	'avocado tree'

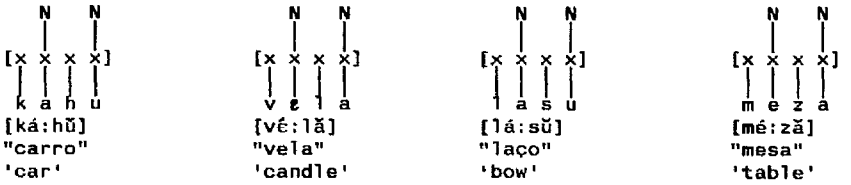
As the examples given in (156) illustrate, affixes may require a specific spell-out for the GS position that follows it, as in the case of -ada-, where the final -a fills the GS position and creates a feminine word. Alternatively, they can maintain a closer link with the stem, in which case the gender of the initial stem is used in the GS position. According to the distinction we made earlier, words ending in the suffix [ada] are feminine and words where the suffix -eir- [ejr] is involved can be either feminine or masculine, depending on the gender of the initial stem. In this strict sense we can say that the gender of these derivatives is predictable, even though it is the suffix which has determined that this be the case.

This means that suffixes may have specific requirements concerning a possible GS that follows it. If we now go back to (155), we will see that if we derive an adjective from a noun which requires no GS, as in a word like [fló:], the fact that the deriving adjective requires a spell-out which follows the general pattern can also apply. This is the case in [flòrí:dŭ]. But an adjective may require a single GS spell-out for both genders as well. This is the case in [flòrístă]. It is important to notice that the derivation of [àlărmá:dŭ] and [flòrí:dŭ] follows the general pattern whereby the feminine is possible, i.e. by adding -a instead of -u in the gender suffix position.

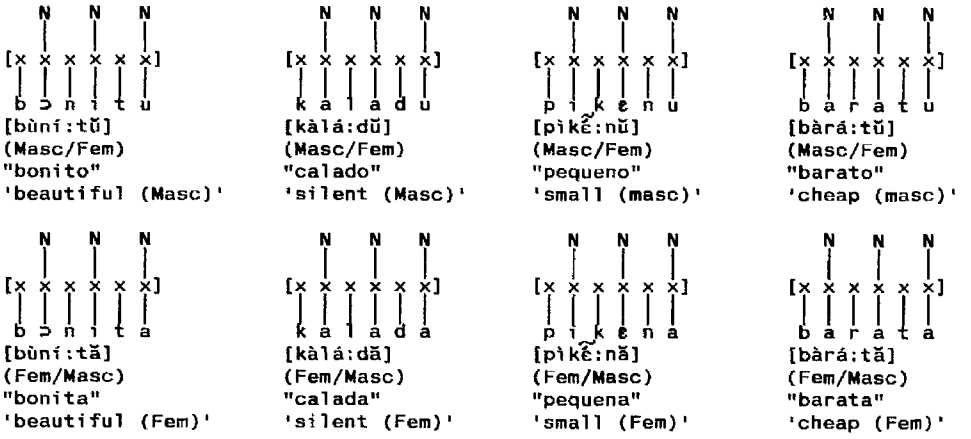
We are now ready for a discussion of the lexical properties of each of these derivations. We have our basic tool which allows us to test the status of a specific string in a derivation, and we now turn to a discussion of the lexical representation of nouns and adjectives as we discuss the status of each derivation.

The simplest of all cases, that is, words whose gender suffix follows the [a] for the feminine and [u] for the masculine pattern, are lexically represented as follows:

(157) a. Nouns



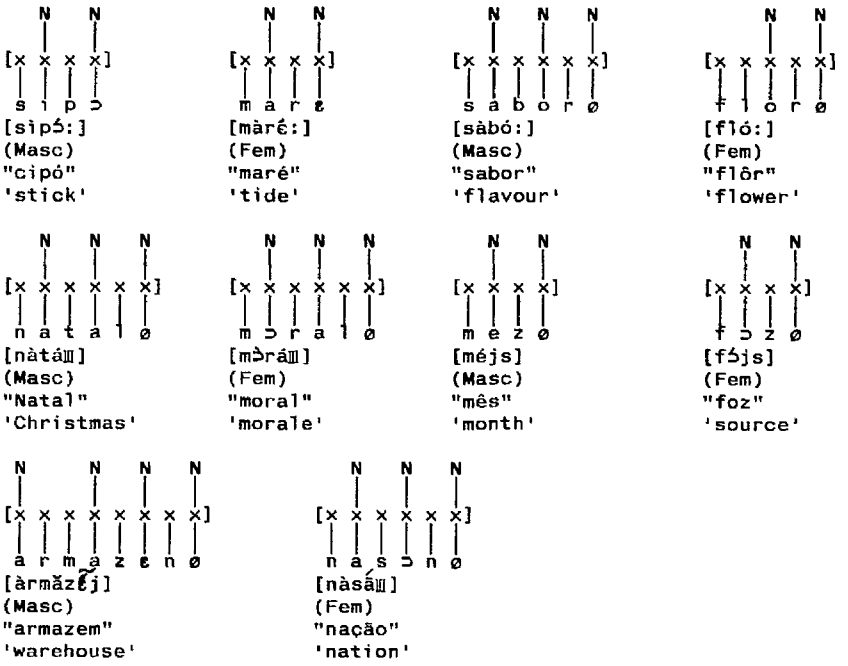
b. Adjectives



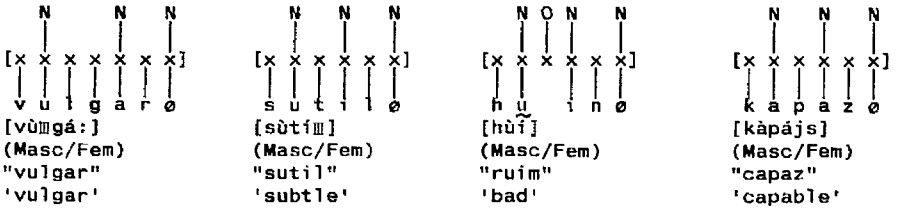
The examples in ((157)a) illustrate cases where the GS position is not open for the distinction masculine/feminine. In ((157)b), we have the lexical representation of adjectives whose GS position is flexible in the sense that the speaker learns either or both of them and applies all the other productive word formation "rules" that he/she has learned for generating other forms.

The next set of cases are those whose gender is not overt, i.e. they have no gender suffix. Because their gender is not explicitly stated like the preceding ones, we are forced to argue that they are specified in the lexicon. In cognitive terms, this means that when speakers learn these words, they also learn their gender separately. It is crucial to understand the nature of the information which it is possible to encode in the lexicon. While it is obvious that certain words have no GS position simply because no segment is realised in this position, we still have to encode some information concerning the gender of these words in their lexical representation, because this property is necessary for the derivation of other words. We propose that these words are lexically represented as follows:

(158) a. Nouns



b. Adjectives



These lexical representations have been proposed taking the following derivations into account:

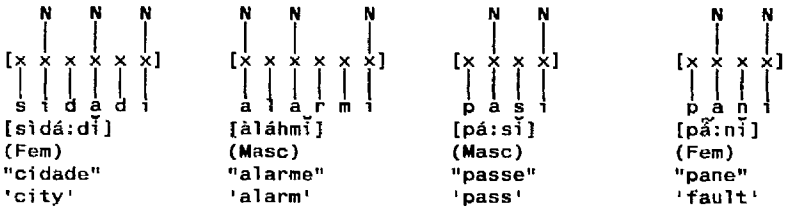
(159)	[sipɔ̆zĩ̃́:ñŭ]	"cipozinho"	'little stick'	(Masc)
	[màrɛ̆zĩ̃́:ñă]	"marezinha"	'little tide'	(Fem)
	[sàbŏró:zŭ]	"saboroso"	'tasteful'	(Masc/Fem)
	[flòrístă]	"florista"	'florist'	(FC)
	[nàtălɛ̃́sĭ]	"natalense"	'native of Natal'	(FC)
	[mòrălístă]	"moralista"	'moralist'	(FC)
	[mézĭs]	"meses"	'months'	(Masc)
	[fɔ́zĭs]	"fozes"	'sources'	(Fem)
	[àrmàzɛ̆nážɛ̃̆j]	"armazenagem"	'storage'	(Fem)
	[nàsjŏnáɯ]	"nacional"	'national'	(FC)[52]
	[vùɯgărĭdá:dĭ]	"vulgaridade"	'vulgarity'	(Fem)
	[sùt̪ĭlé:ză]	"sutileza"	'subtleness'	(Fem)
	[hùĩdá:dĭ]	"ruindade"	'badness'	(Fem)
	[kàpàsĭdá:dĭ]	"capacidade"	'capacity'	(Fem)

[52]Since our aim in this discussion is to prove the integrity of the stem, in this discussion we will ignore the appearance of [j] in the stem nucleus.

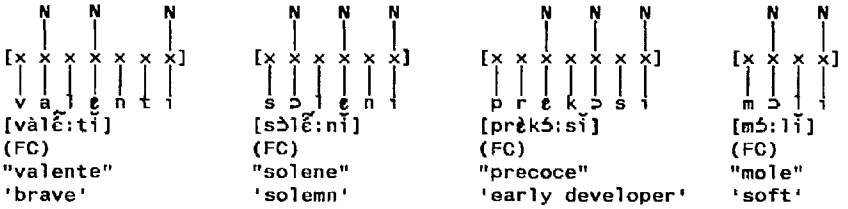
The lexical representations proposed in (157) and (158) represent what we believe to be the only way to include all the information necessary for the derivation of other forms. Further research might prove that there are alternative ways to represent these forms lexically, but at the present stage of our research this seems less than obvious. We have to admit that it may be necessary for other morphological or even phonological information to be encoded in them, but this does not follow from this discussion.

The final set of words for which we would like to propose a lexical representation includes words whose GS position is filled with [i] "-e". In these cases, the speaker has to learn the gender in each case because he/she also learns that the [i] that surfaces in the GS position does not follow the regular pattern.

(160) a. Nouns



b. Adjectives



The spell-out of the gender suffix as either -ø or -i deserves some discussion. Although we will not go into a full analysis of the licensing of domain final empty nuclei in BP, we believe that

the realisation of some of the gender suffixes presented here is certainly affected by this parameter. We believe that word final empty nuclei are not licensed in BP. This of course predicts that, since they are unlicensed, word final empty nuclei should be interpreted in some way. We believe that the [i] -i type of spell-out of the GS represents the interpretation of unlicensed word final empty nuclei in BP. This means that we should not expect to find uninterpreted word final empty nuclei in BP. However, certain derivations seem to still end up with final empty nuclei which are not interpreted. In nouns and adjectives, these cases are limited to those whose stem ends in one of four consonants: -n-, -r-, -l-, or -s-. Similar cases occur in adverbs and other grammatical categories, as shown below:

(161) a. Adverbs

[máɪ]	"mal"	'badly'
[dimájs]	"demais"	'too much'

b. Prepositions

[májs]	"mas"	'but'

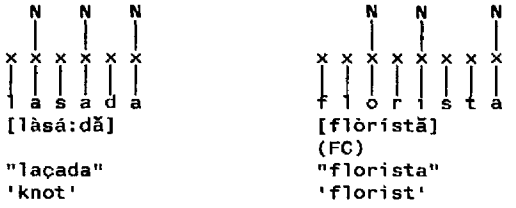
c. Conjunctions

[sẽj]	"sem"	'without'
[kɔ̃:]	"com"	'with'

However, this does not mean that there are no cases of [i] as the spell-out of the gender suffix following one of these consonants. There are a number of cases where -n-, -l-, -r-, and -s- are followed by [i], as given in (160). The gender suffix in some cases, however, reaches the end of the derivation still uninterpreted. It is exactly in these cases that diphthongs are derived. For this reason, we believe that the fact that they remain empty until the end of the derivation is responsible for the processes that take place in these derivations and which result in the formation of diphthongs. The reason why these processes occur only when -s-, -n-, -r- or -l- are involved needs further investigation before an adequate explanation can be reached. The reader is referred to Da Silva (1992) for a full analysis of the derivation of diphthongs and a discussion of empty nuclei in word final position in BP.

The lexical representation of words whose derivation includes one or more suffixes is illustrated below:

(162) a. Nouns

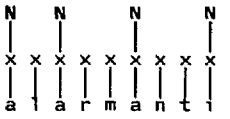


[làsá:dă]

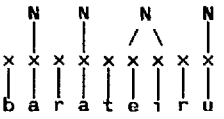
"laçada"
'knot'

[flòrístă]
(FC)
"florista"
'florist'

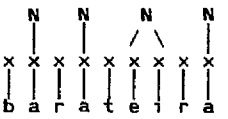
b. Adjectives



(FC)
[àlãhmãtĩ]
"alarmante"
'alarming'



(Masc/Fem)
[bàrătéjrŭ]
"barateiro"
'man who sells items cheaply'



(Fem/Masc)
[bàrătéjră]
"barateira"
'woman who sells items cheaply'

Just as we saw for the cases where no suffixes are involved, in some cases the spell-out of the gender is predictable. Although both forms are present lexically, the speaker can easily generate the other gender based both on the extra information concerning the gender and on the word formation rules that he/she acquires.

In this section we have analysed the morphology of nouns and adjectives in BP. We have proposed that these two categories share the same morphological structure, namely a stem followed by a gender suffix, but that certain words deviate from this morphological structure in that they lack a gender suffix. We have proposed that, whenever present, the gender suffix is always found at the end of a domain.

In the next section we analyse the assignment of stress to these two grammatical categories based on the morphological structure presented in this section.

6.3 *The Assignment of Stress to Nouns and Adjectives in BP.*

In the previous section, we discussed the morphology of nouns and adjectives in BP. We proposed that these two categories share the same general morphological structure whereby a stem, in the majority of cases, is followed by a gender suffix.

In this section we will concentrate on how stress is assigned to words belonging to these two categories. We will attempt to prove that nouns and adjectives are assigned stress with the same set of parameters used for verbs. The metrical settings proposed are repeated below:

(163) i. Feet are binary and left-headed.
ii. Foot construction begins at the right edge of the word.
iii. Word trees are right-headed.
iv. BP is quantity insensitive.

Before we can consider the actual data, it is important to note that one set of words must be included in the analysis of nouns and adjectives: the infinitives. The infinitives are illustrated below:

(164)			
	[fàlá:]	"falar"	'to speak'
	[bèbé:]	"beber"	'to drink'
	[fìrí:]	"ferir"	'to hurt'

Although we know that infinitives are morphologically composed of a stem + thematic vowel (TV) + the nominalizer -rø-, and that their internal morphological structure dictates that they should be treated like verbs, we would like to propose that these structures have a dual behaviour in that they display characteristics of both nouns and adjectives, and of verbs. First of all, we propose that their internal morphological structure is phonologically irrelevant. We have good reason to believe they should receive the same treatment as nouns and adjectives. There is some very simple morphological evidence to support the interpretation of infinitives as nouns. Specifically, these forms can be pluralised, as shown below:

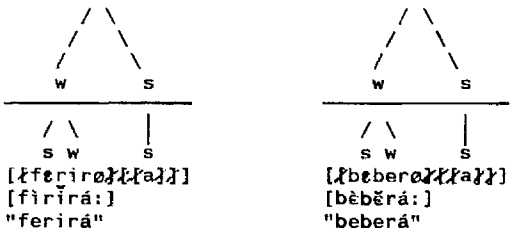
(165) "Os falares do Nordeste identificam as origens de cada povo"

'The languages of Northeast Brazil identify the origins of the each group of people'

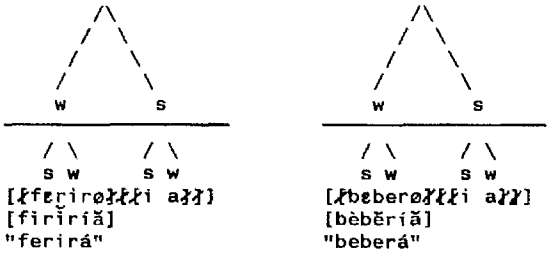
On the other hand, the fact that they participate in the formation of the future and the conditional suggests a pattern of behaviour belonging to verbs. We have seen in chapter 3 that under those circumstances they are stressed like any other verb form.

By proposing that infinitives are treated like nouns and adjectives, we have to explain not only how they are stressed while behaving like nominal categories but also how they could be treated differently by the phonology. We have seen in chapter 3 that when infinitives are used in finite forms, for instance in the derivation of the future and the conditional, they are stressed exactly like the remainder of the finite verb forms. The proposed metrical settings apply, and yield the following stress pattern:

(166) a. The Future.



b. The Conditional.



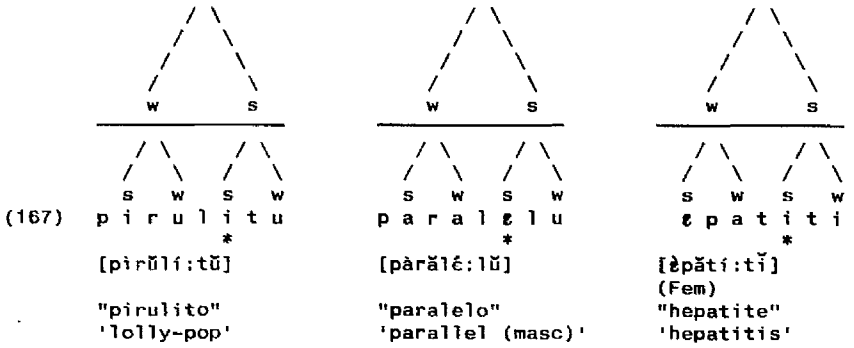
We propose that the phonology should in some way be able to see when a given form is being used as a noun, rather than as a verb form. Other aspects, such as the fact that infinitives undergo vowel harmony processes exactly like all the other verb forms, would have to be transparent to the phonology in favour of their interpretation as verbs. Further research must be

carried out before we can see clearly how this double treatment would be encoded in the lexical representations. These views concerning the infinitive carry a considerable significance in theoretical terms. The Minimalist Hypothesis, according to which processes apply when the conditions that trigger them are satisfied, proposed in Government Phonology (Kaye, 1992), seems to be too restrictive. Although this topic deserves much more investigation than we have already carried out in this respect, the present results suggest that the phonology should recognise grammatically conditioned generalisations in some way.

We now turn to the discussion of how stress is assigned in nouns and adjectives.

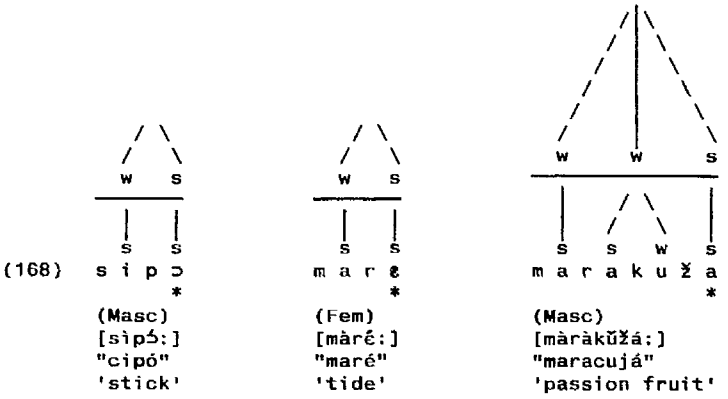
Let us consider the basic type of structure in these categories, that is, cases where the gender suffix falls under the general pattern of [a] -a for the feminine and [u] -u for the masculine and no affixes are involved. The data analysed thus far have shown us that primary stress always falls on the final nucleus of the stem. Starting with the basic cases, let us suppose that the stem final nucleus carries a lexical marking[53]. Let us also suppose that this lexical marking has to coincide with the final strong node of the word in metrical structure. Let us apply the metrical settings proposed for BP to a few words with the simplest morphological structure. This will result in the following metrical analysis:

[53]Garde (1973) believes that roots are stressed on their final syllable or their penult in all Romance languages, as shown in the Italian words "mattín-a", "cittá", and "óper-a", except in Provençal and in French, and that this stress is specified in the lexicon of the language.



As the realisations in square brackets illustrate, the metrical settings proposed earlier for the assignment of stress to verbs can correctly assign stress to these forms. The metrical settings have applied to the examples above and a foot has been constructed in such a way that the lexically marked nucleus coincides with the final strong node of the word. The final nucleus, the gender suffix, has been assigned the role of weak member of the final foot.

Let us now extend this hypothesis to the cases where the GS does not surface and the stem ends in a nucleus.

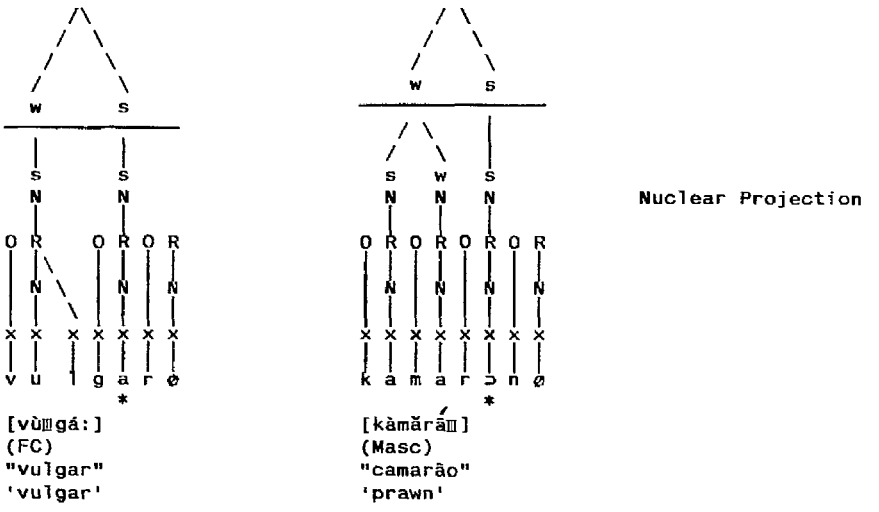


As illustrated above, the metrical structure constructed on words which lack a gender suffix predicts the correct pattern of stress. In these words, the lexical marking also coincides with the final nucleus of the stem and a degenerate foot is constructed on this nucleus.

Let us now apply this hypothesis to cases whose stem ends in a

consonant but the GS position remains empty throughout the derivation[54].

(169)



As illustrated above, the construction of metrical structure of words which have no gender suffix predicts the correct pattern of stress, both in the cases where the stem ends in a nucleus and when it ends in a consonant. According to Government Phonology, the syllabification of the word final consonant with the application of the Coda Licensing and the Domain Final Empty Nucleus Licensing Principles, predicts the presence of an empty nucleus at the end of the words in (169). The Coda Licensing principle requires that a consonant in a rhymal position be governed by a following onset position. This principle rules out the possibility of analysing a final consonant as being in a rhymal position, because it is not followed by an onset. And if a final consonant is not in a rhymal position, the only remaining non-nuclear position it can be is the onset. The Licensing principle will predict that this final onset is followed by a nucleus, which in these cases is always empty. The Licensing Principle is repeated below:

---

[54]In the following derivations, for typographical reasons as well as not to distract the reader with matters which are irrelevant to the discussion, we will ignore the derivation of the diphthong and of the long vowel.

(170) Licensing Principle (Kaye (1990b))

All phonological positions save one must be licensed within a domain. The unlicensed position is the head of the domain.

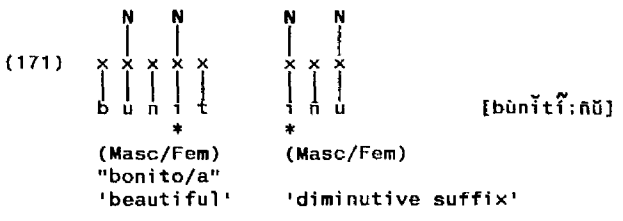
This being the case, the theory predicts that an empty nucleus follows the stem final consonant in (169). Let us now return to the assignment of stress to these forms. Since stress is assigned at the nuclear projection and the domain final empty nucleus is not present at this projection, a degenerate foot is constructed on the stem final (lexically marked) nucleus.

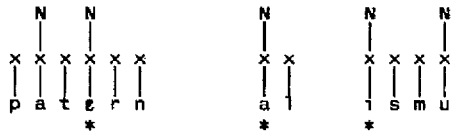
So far we have analysed cases where the morphological structure is stem+GS. We have also seen how stress is assigned to cases where the GS is absent. In these cases a degenerate foot is constructed on the stem final nucleus which is lexically marked. Let us now consider cases where affixes are involved. In the same way that we proposed that the final nucleus of the stem carries a lexical marking, we will propose that affixes as well are lexically marked on their final nucleus, which excludes the GS position. The lexical marking in affixes will also mean that it has to coincide with the rightmost strong node in the construction of metrical structure.

We have now reached a conflicting situation in the derivation. Whenever affixes are involved, a sequence of lexical markings will arise, namely, one in the stem and one for each affix involved in the derivation.

This conflicting situation is shown below:

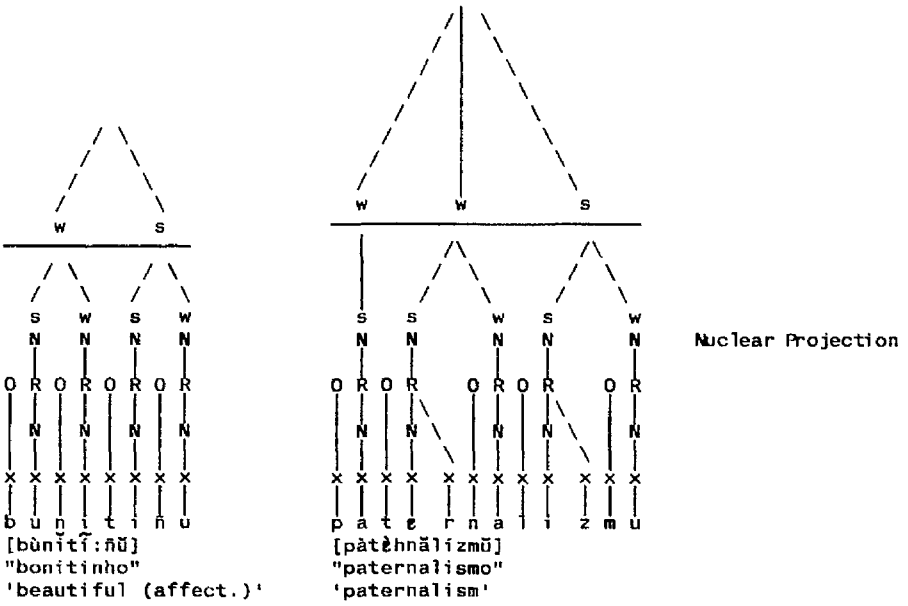
(171)





As illustrated above, derivations where affixes are involved result in a sequence of lexically marked nuclei. We propose that this problematic situation is resolved by the application of a parameter whereby the final affix is interpreted as dominant. This means that, phonologically, the final lexical marking will override others preceding it, and the final strong node has to coincide with it in the construction of metrical structure. Let us now apply the metrical settings proposed for BP to these forms:

(172)

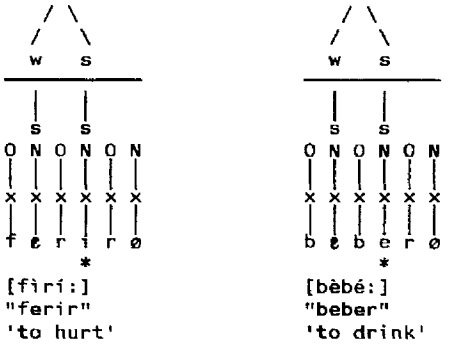


As seen above, the proposal that the final lexical marking indicates the place where the final strong node will be predicts the correct results. It is important to keep in mind that we have morphological support for this proposal, namely the fact that the final suffix determines the type of spell-out the newly derived word will have.

If we now re-consider the infinitives, according to this analysis

we will have to propose that they are also lexically marked on the thematic vowel. While behaving like nouns, the internal morphological structure of the infinitives is irrelevant. The sequence verb stem+thematic vowel+nominalizer will behave exactly like any of the stems or derivational stems we have considered in this chapter. The metrical analysis where one such lexical marking is taken into account is illustrated below:

(173)



The metrical analysis above predicts the correct pattern of stress for the infinitives. Exactly as described for the examples in (169), the theory predicts the presence of an empty nucleus at the end of the infinitives. This final empty nucleus will not be taken into account in the construction of metrical structure because it is not present at the projection where metrical structure is constructed. Further research needs to be carried out in this respect, but the present results suggest that the phonology should recognise grammatically conditioned generalisations in some way.

This concludes our discussion on the assignment of stress to the nominal categories in BP. We have proposed that they are stressed with the same set of parameters used for the assignment of stress to verbs. We have seen, however, that in the assignment of stress to nominal categories the final strong node has to coincide with a lexical marking which indicates the final nucleus of the (derivational) stem. The analysis has shown that, regardless of the number of affixes involved in a derivation, the final nucleus of the final affix (excluding the GS) will receive primary stress.

In this chapter we have analysed the derivation of nouns and adjectives whose morphology involves non-analytic domains only. In the next section we will examine forms whose internal morphological structure is relevant from a phonological point of view.

7.1 *Introduction.*

The nouns and adjectives whose metrical structure was analysed in the preceding chapter present the type of morphological structure to which the phonology is not sensitive. In this chapter we will examine the assignment of stress to nouns and adjectives whose internal morphological structure is relevant from a phonological point of view. We will offer an analysis of stress clash and its resolutions.

7.2 *Basic Assumptions and Affixes in BP.*

In the analysis of nouns and adjectives to be presented in this section, we will maintain the view defended in chapter 6 according to which the GS is domain-final in nouns and adjectives in BP. We saw that derivations where non-analytic suffixes are involved display only one GS position, i.e. domain finally, unless it is not present in a given form. Since we argue that one GS position may be found at the end of each domain, it follows that we argue that there are as many gender suffixes as the number of domains involved in a derivation in BP, unless of course a GS is not present in the derivation which includes a specific stem. According to Government Phonology, analytic suffixes of type ((174)b.ii) are always preceded by a domain boundary. If we consider that in the majority of cases a GS is required (domain-finally), we can conclude that in BP analytic suffixes are in the majority of cases preceded by a GS position. (174) illustrates the case where two strings A and B are involved and GS positions are required by both A and B.

(174) a. Non-Analytic Derivation.

[A + B + GS]

b. Analytic Derivation.

i. 1. * [A [B + GS]]

2. [[A + GS] B]

ii. [[A + GS] [B + GS]]

In (174), a GS position is added at the end of each domain. Although the structures in ((174)b.i) are logically possible, the structure in ((174)b.i.1) does not appear to be attested (Kaye, to appear), and we will argue that the structure in ((174)b.i.2) does not exist in the derivation of nouns and adjectives in BP. This means that in this analysis we will deal with the structures in ((174)a) and ((174)b.ii) only.

It is crucial to keep in mind the claims made in the preceding chapter with respect to the nature of the morphology of nouns and adjectives in BP, such as the flexibility of the GS in non-analytic derivations. It is important to bear in mind that they are affixed only in the final stage of a derivation. The lexical gender information which is found lexically will then surface. These claims are, of course, maintained throughout this analysis.

From a theoretical point of view, one further point needs to be made clear. Once stress is assigned it cannot be removed. There is no deforestation. It is important to contrast these assumptions with those proposed in other contemporary frameworks such as Halle & Vergnaud (1987). In their framework, stress assigned in a "cyclic" domain by the English Main Stress Rule is wiped out in the next cycle of the derivation (Halle & Vergnaud, 1987; 81). According to them, the phonological component is divided into at least two strata which are termed the cyclic phonology and the noncyclic phonology. Moreover, the principle of Strict Cyclicity governs the application of rules in the cyclic strata only. In our framework, processes are subject to the Minimalist Hypothesis, which is given below.

(175) Processes apply whenever the conditions that trigger them are satisfied (Kaye (1992)).

According to this hypothesis, no matter what stage of a derivation has been reached, if the context for the application of a particular process is met, that process will take place. Furthermore, processes are invariably subject to the principle of Strict Cyclicity (Kean, 1974). This means that after the deletion of

boundaries, the result of the application of a certain process such as stress assignment, for example, cannot be changed. Our position in this respect gives completely different results from those of Halle & Vergnaud, as illustrated in the preceding paragraph. We now turn to actual cases of affixation and examine a number of affixes in BP.

7.3 *The Derivation of Nouns and Adjectives in BP.*

We will start by examining some non-analytic cases displaying the ((174)a) type of derivation, before we can move on to derivations where the ((174)b.ii) type of derivation is observed. The discussion presented above and in the preceding chapter has allowed us to conclude that non-analytic suffixes should be attached directly to the stem or to a derivational stem before a GS position can be added, if such is the case. Once again, we have to make clear that, by arguing that one suffix is non-analytic or analytic, we must be aware of the implications of such a statement in BP. Since we assume that the GS position is domain final in BP, a suffix **y** which we claim to be analytic will always be preceded by a domain boundary. Since a domain boundary implies the presence of a preceding domain, an analytic suffix will be preceded by a GS position if the preceding domain ends in a GS. Other than language-particular issues, there are theoretical aspects which need to be kept in mind. For example, the presence of a domain boundary will always entail the application of the domain final Licensing of empty nuclei and the Coda Licensing principles, as well as the application of any other process which is sensitive to domain final position.

We propose that the suffixes in (176) below are non-analytic.

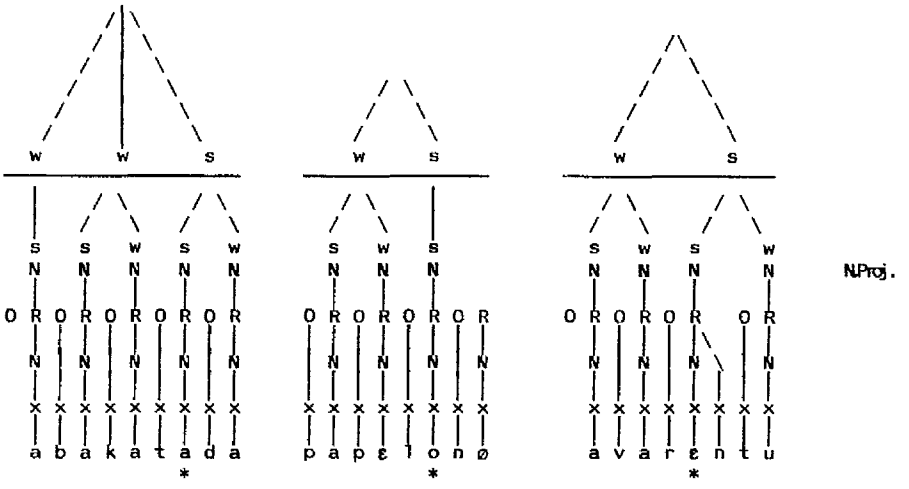
(176)	-ada-	-ɔnø-	-ɛntu-
	*	*	*
	-ista-	-eiru-	-udu-
	*	*	*
	-ensi-	-ɔzu-	-alø-
	*	*	*

If we assume that the suffixes in (176) are non-analytic, they will obligatorily have to be attached to the preceding stem or derivational stem in a derivation. Their GS will be realised if they are domain-final. We must now apply the metrical settings proposed for BP and check whether this produces the attested outcome. The metrical settings are repeated below:

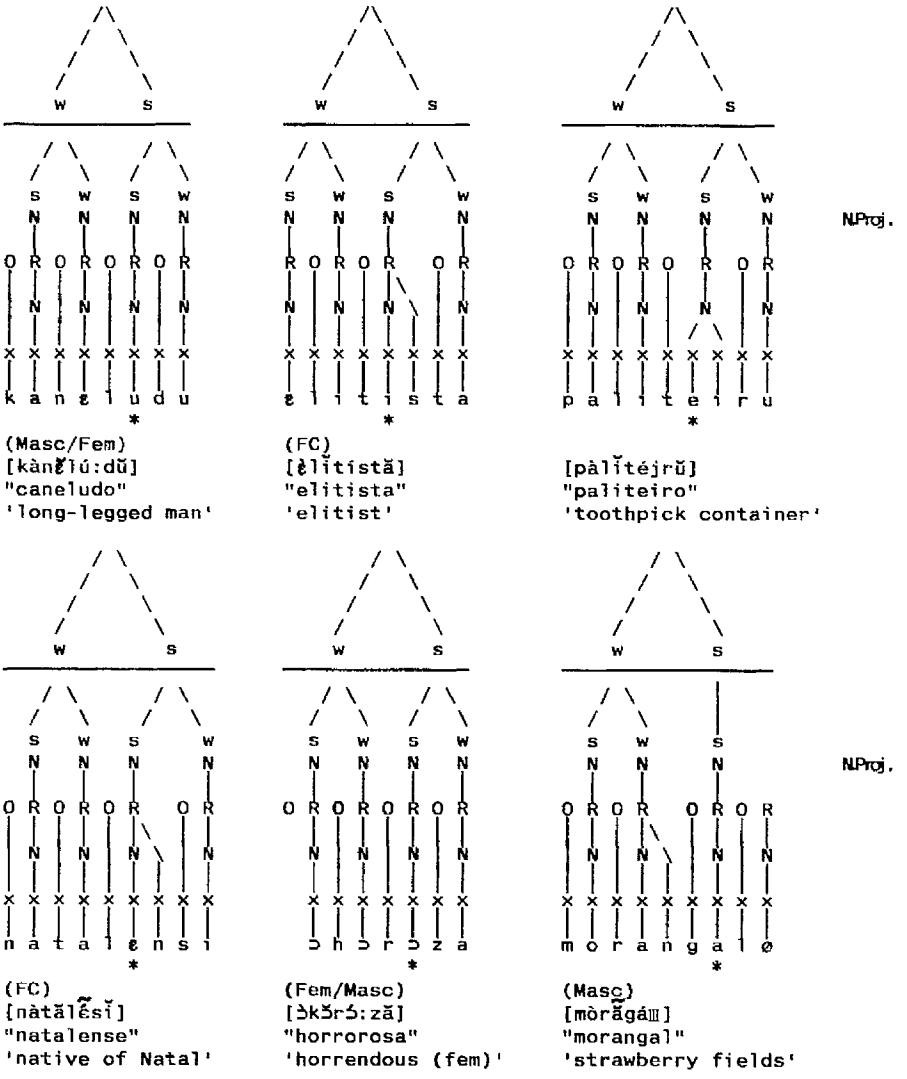
(177) i. Feet are binary and left-headed.
ii. Foot construction begins at the right edge of the word.
iii. Word trees are right-headed.
iv. BP is quantity insensitive.

This analysis correctly predicts the attested forms, as the examples below prove[55].

(178)



[55]In the non-analytic derivations that follow only the lexical marking and gender information concerning the gender of the suffix at issue have been given.



In each of the above cases, one non-analytic suffix is attached to the stem. The GS position surfaces according to the final suffix. If this analysis is correct, stress should be assigned to the whole string in one single application. The attested pattern of stress given in each case proves that the interpretation of the suffixes -ada-, -onø-, -entu-, -ista-, -udu-, -ensi-, -ɔza-, -eiru- and -alø- as non-analytic is correct. As they are the final suffix in the derivation, their GS is realised, and their lexical marking indicates the nucleus which has to coincide with the final strong node. Stress will then be assigned to the entire string as if no morphological boundaries existed, as in -paralɛlɛpípɛd+u- [pàrălɛ̀lɛ̆pípɛ̆dŭ] "paralelepípedo" 'curb'.

Let us now consider the following data.

(179)	[kàfɛ̌zá͡ɯ]	"cafezal"	'coffee fields'
	[màrɛ̌zɔ̂:nă]	"marezona"	'big tide'
	[àbàkăšìzá͡ɯ]	"abacaxizal"	'pineapple fields'
	[nàzărɛ̀zêsì]	"Nazarezense"	'native of Nazaré'
	[màràkŭžàzéjrŭ]	"maracujazeiro"	'passion fruit tree'

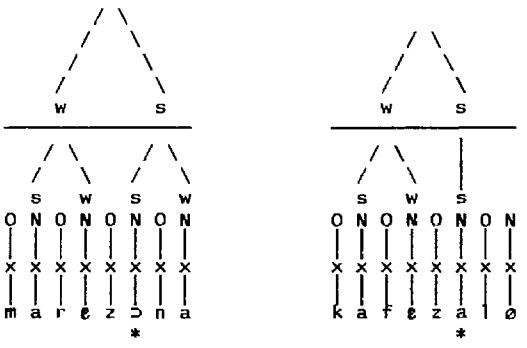
Because we have seen these stems in other derivations earlier in the discussion, we can easily carry out the morphological parsing of these forms, as shown below:

(180)	[kàfɛ̌ + zá͡ɯ]	"cafezal"	'coffee plantation'
	[màrɛ̌ + zɔ̂:nă]	"marezona"	'big tide'
	[àbàkăšì + zá͡ɯ]	"abacaxizal"	'pineapple fields'
	[nàzărɛ̀ + zêsì]	"Nazarezense"	'native of Nazaré'
	[màràkŭžà + zéjrŭ]	"maracujazeiro"	'passion fruit tree'

As illustrated above, we are proposing that the words in (180) derive from "café", "maré", "abacaxi", "Nazaré" and "maracujá" by adding "-zal", "-zona", "-zense", and "-zeiro".

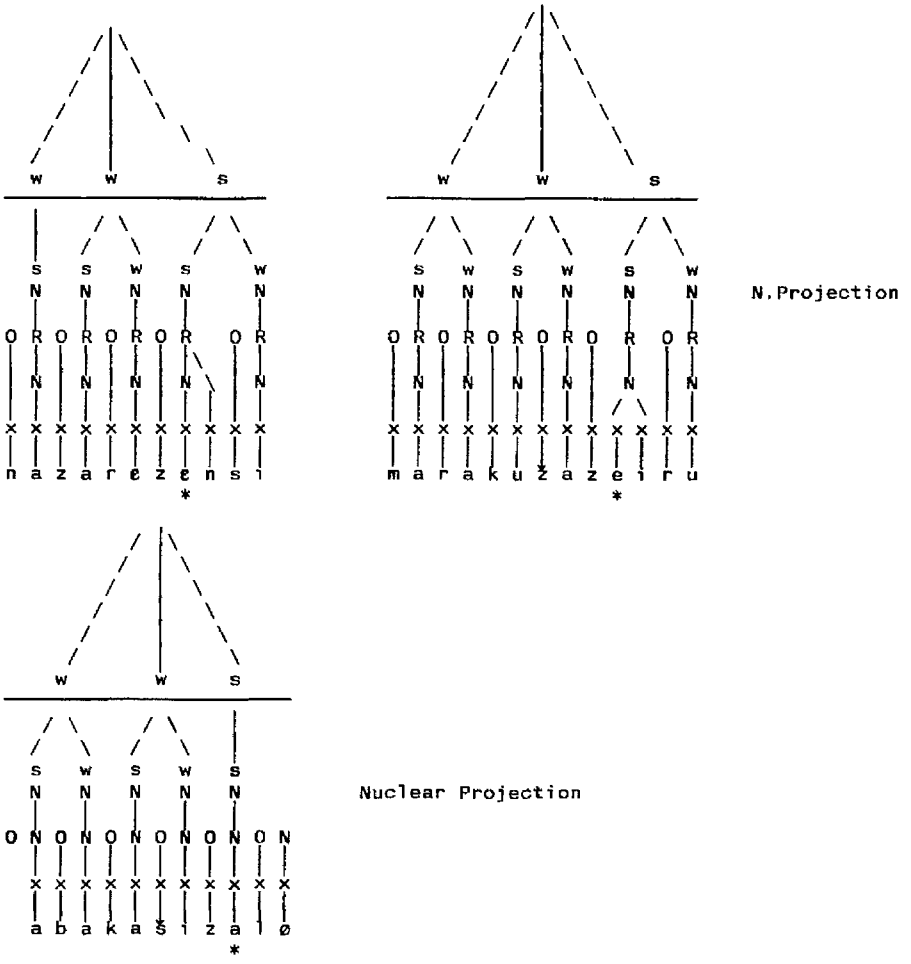
However, the question arises as to how these forms are to be assigned stress. Let us take the shortest words. If we apply the metrical settings proposed thus far to the whole string as if we were dealing with a non-analytic derivation, then the following metrical structures would result:

(181)



The metrical structures proposed in (181) predict the attested pattern of stress for these forms, as given in (180). Let us now apply the metrical settings to longer words, such as the ones below:

(182)



According to the attested pattern for these words given in (179), the metrical structures constructed on the strings above predict an incorrect pattern of stress. The metrical structures above predict the following stress pattern:

(183) * [nàzàrž̃zé̞sǐ]
* [màrākùžāzéjrŭ]
* [àbākàšǐzáɯ]

We have seen that the analysis proposed so far predicts the correct pattern of stress for words whose stem, before the suffix is attached, consists maximally of two nuclei. Words whose stem consists of three or more nuclei cannot be accounted for by this analysis.

According to Kaye (1992, PC), although affixes are usually stable

with respect to their analytic/non-analytic status, it has been shown that some suffixes can display special behaviour which cannot be determined on a binary basis. This special type of suffix arises in English, for example, where suffixes like -metre-, -able- and -ible- are known to be either analytic or non-analytic, depending on each individual case. This means that their status cannot be defined independently of the derivation. The examples below illustrate derivations where these suffixes are involved.

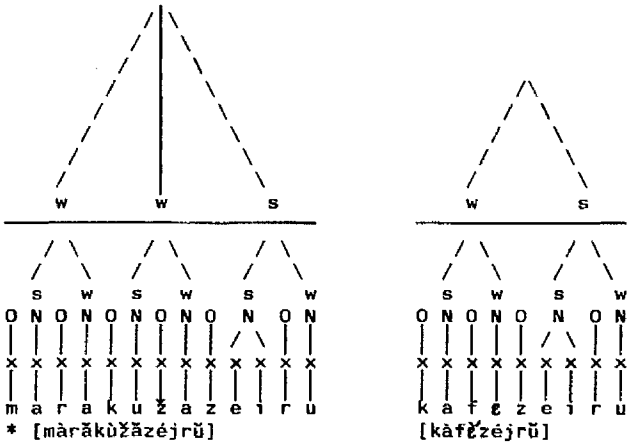
(184) [thermómeter]
[[cénti][mètre]]

Let us now return to the BP data. In order not to have to explain the derivations in (179) as cases where suffixes display a different status with respect to analyticity according to the derivation at issue, as shown in (184), we have to test the behaviour of these suffixes when they are combined with all possible types of stems. One simple test we can do is to combine the various suffixes with the various stems involved in the derivations illustrated in (179). This is done below[56]:

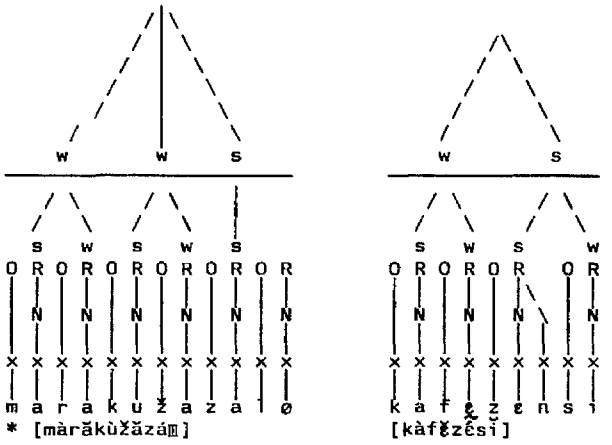
(185)	[màràkŭžàzéjrŭ]	"maracujazeiro"	'passion fruit tree'
	[kàfɛ̆zéjrŭ]	"cafezeiro"	'coffee tree (H)'
	[màràkŭžàzáɯ]	"maracujazal"	'passion fruit fields'
	[kàfɛ̆zɛ̂sĭ]	"cafezense"	'native of Café (H)'

The metrical analysis of these forms is given below:

(186)



[56](H) stands for hypothetical words.



As shown above, again the combinations where long stems are involved invariably predict an incorrect pattern of stress.

Although further research might prove that this analysis is not correct and that there is an alternative explanation for these facts, here we will offer the following provisional analysis.

We would like to propose that the data in (179) involve analytic derivations. We will begin with derivations where longer stems are involved, and, once we have achieved a successful analysis of these derivations, we will then consider shorter stems such as "café-" and "maré-".

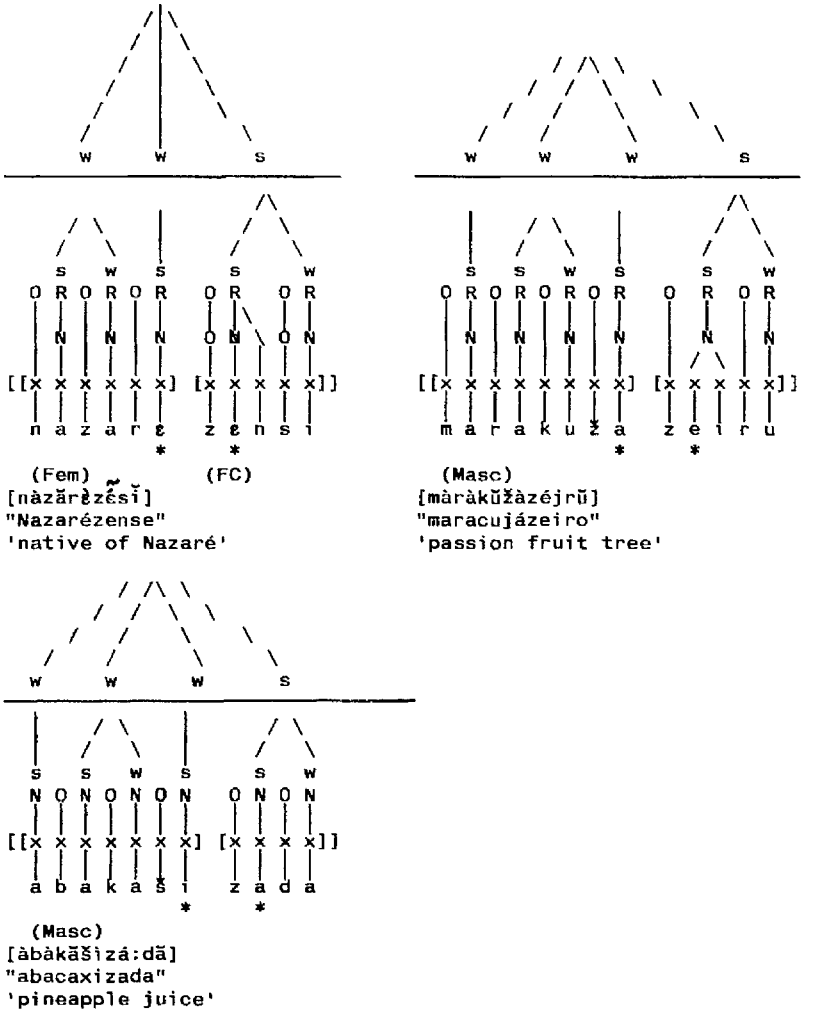
We propose that the morphological parsing for the words in (179) is as follows:

(187)	[[nàzărɛ̃][zɛ̃sĭ]] * * (FC)	"Nazarézense"	'native of Nazaré'
	[[màràkŭžà][zéjrũ]] * * (Masc)	"maracujázeiro"	'passion fruit tree'
	[[àbàkăši][zá:dă]] * * (Fem)	"abacaxizada"	'pineapple juice'

Stems, which in this case are characterized by the fact that they do not have a GS, constitute one domain, and are followed by one analytic suffix which has a GS in another domain. In these structures, each domain behaves in an independent fashion. The expected metrical analysis is one where stress is assigned to each

individual domain independently. Since the leftmost domains have no GS position, stress will be assigned in such a way that the final (lexically marked) nucleus coincides with the final strong node of the domain. Stress is, of course, assigned independently to the rightmost domain. In this latter domain, the GS will be assigned the role of weak member of the foot. Once stress is assigned to each individual domain, when the derivation reaches the larger domain where the individual domain boundaries are deleted, the word tree is constructed on the projected heads (of previous domains whose boundaries have been deleted). Notice that the word tree is only constructed after the deletion of each individual domain boundary, when the word (=larger) domain is reached. The metrical analysis of the words in (187) is given below.

(188)



As the examples above show, the metrical structure constructed in such a fashion predicts the correct pattern of stress. In these examples, the string consists of two individual domains where processes apply independently.

In order to test the reliability of this analysis we need to give independent evidence for the analytic status of these suffixes. It is clear that the pattern of stress predicted by this analysis is correct. One such piece of evidence would be an example where a GS is present and clearly audible in the leftmost domain. In this respect, the only evidence for this is the derivation of the plural in these forms.

Before we can understand how the plural of these forms can help us in this respect, we need to make clear how plurals are realised. Let us consider the forms without the suffixes which are at issue here. They are given below:

(189)	[nàzăré:]	"Nazaré"	'Nazaré - town'
	*		
	(Fem)		
	[màràkŭžá:]	"maracujá"	'passion fruit'
	*		
	(Masc)		
	[àbàkăší:]	"abacaxi"	'pineapple'
	*		
	(Masc)		

We now proceed to the evidence as follows. We argue that the plural is always found domain-finally. Derivations where the plural is involved are illustrated below:

(190)	SINGULAR	PLURAL		
	[ká:zã]	[ká:zăs]	"casa(s)"	'house(s)'
	[kàzĩ:ñă]	[kàzĩ:ñăs]	"casinha(s)"	'little house(s)'
	[kàzõ:nă]	[kàzõ:năs]	"casona(s)"	'big house(s)'

Let us now pluralise the forms in (189). This is done below:

(191)	[nàzărɛ́js]	"Nazarés"	'Nazarés - towns'
	[màràkŭžájs]	"maracujás"	'passion fruits'
	[àbàkăšijs]	"abacaxis"	'pineapples'

Without entering into the details of the derivation of the plural in

these forms, it is important to point out that the realisations given above are those from Natal. A simple rule to determine when a diphthong is derived in the plural is to say that they are derived when gender suffixes are absent. Let us now go back to the cases in (187). In order to determine whether there is a domain boundary between the stem and the suffix, we simply have to pluralise these forms. If a diphthong is derived after the initial stem, it means that there is a plural suffix there, i.e. domain finally. Otherwise, no domain boundary exists and the suffixes have to be analysed as non-analytic. Their plurals are given below:

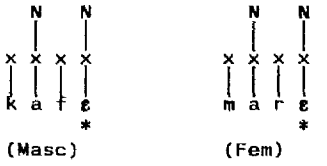
(192)	[nàzărɐ̀jzɛ̃́sɨ̌s]	"nazarezenses"	'natives of Nazaré'
	[màràkŭ̌ʒàjzéjrŭs]	"maracujázeiros"	'passion fruit trees'
	[àbàkăšɨ̀jzá:dăs]	"abacaxizadas"	'pineapple juices'

We now have independent evidence to support our proposal that these suffixes are analytic. To recapitulate, in order to prove that the "-zal", "-zona", "-zense", and "-zeiro" suffixes are analytic we used forms where the plural, which we claim to be domain-final, is used. Since the plural of forms with no GS results in the derivation of a diphthong in Natal, we pluralised forms where the relevant suffixes are involved to check whether a plural suffix is present by the occurrence of a diphthong. If it were present, a diphthong should be derived between the initial stem and the suffix, otherwise no diphthong would result. The examples in (192) proved that the plural suffix is indeed present between the two strings. We conclude that the "-zal", "-zona", "-zense", and "-zeiro" suffixes are analytic. Once we have evidence for the analytic status of these suffixes, we can now go back to the matter in hand, i.e. how stress is assigned to forms where the stem consists of two nuclei, which is the topic of the following section.

7.4 *Clashes and their Resolution in BP.*

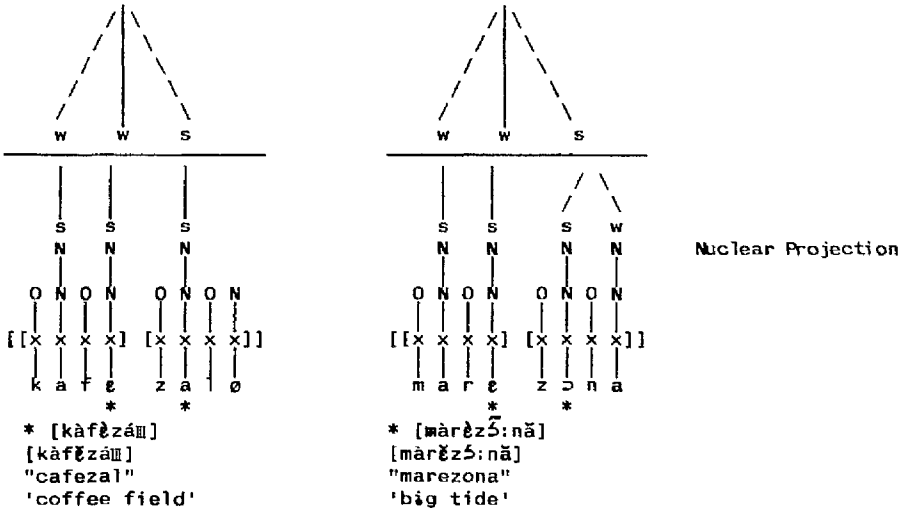
In the preceding section, one type of analytic derivation was intentionally avoided: the case where the stem has two nuclei only and has no gender suffix, as in -kafɛ-, and -sipɔ-. Their lexical representation is given in (193) below.

(193)



We will deal here specifically with derivations where this type of stem is involved. In the previous section we saw how stress is assigned in analytic derivations. In those cases, stress is assigned to each individual domain and, at the deletion of the innermost domain boundaries, no stress is removed. After the deletion of these boundaries, the word tree is constructed (188). Let us now examine cases where the above type of stem is involved.

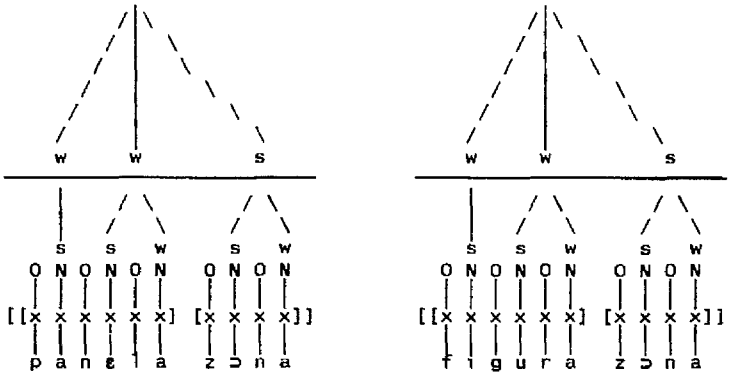
(194)



As the realisations in brackets illustrate, the metrical analysis achieved by applying the set of parameters proposed in chapter 2 for BP is not correct. According to those parameters proposed in that chapter, these forms should display the pattern [ò ò ó] and [ò ó ó ǒ], respectively.

Since the derivation of [[kafɛ][zalø]] and of [[marɛ][zɔna]] are cases where the initial stem requires no GS, it seems reasonable to contrast these cases with examples where the stem has the same number of nuclei but where a GS is present. That is the case of "panelazona", whose metrical structure is constructed on five nuclei[57].

(195)



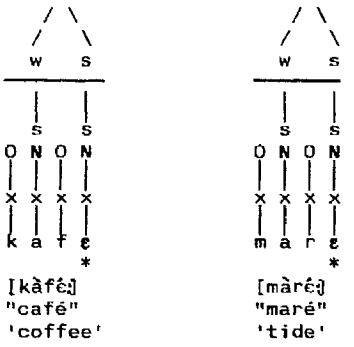
As suggested by the derivations where a GS is present (195) as opposed to those where no GS is present (194), the absence of the GS seems to be responsible for the difference in the stress pattern. In metrical terms, the presence of one weak node between the two strong ones (on the lexically marked nuclei) entails the assignment of a correct pattern of stress.

We therefore propose that the reason behind the incorrect pattern of the structures in (194) is, in some way, linked to the absence of the GS position in "café" and "maré". Since metrical settings can predict the correct pattern of stress for -kafɛ- [kàfɛ́] and -marɛ- [màrɛ́], we can safely conclude that the presence of -zal- and -zona-, respectively, are triggering the shift of stress in

[57]The reader will notice that in the following type of derivation the -zon- suffix only is used. This is due to the fact that it is the only suffix (other than the diminutive, discussed in the following chapter) whose concatenation with stems which have no GS is still acceptable for the Natal speaker. Analytic derivations are, in principle, unacceptable to the Natal speaker if non-analytic derivations are possible.

-kafɛ- and in -marɛ-, respectively. If this were not the reason, then stress shift would occur in the derivations of -kafɛ- 'coffee' and -marɛ- 'tide'. Their metrical analyses below prove that that is not the case.

(196)



It is important to notice that metrically the presence of -zal- and of -zona-, respectively, in (194), means the juxtaposition of one strong node next to the already existing sequence of two strong nodes, triggering the shift of stress. Stress shift has been detected in many languages when two stresses clash, particularly when they are adjacent. As mentioned above, since this phenomenon only occurs in analytic derivations, as illustrated by the examples seen so far, we can safely conclude that it is stress shift that is operative in these cases. As illustrated in (194), metrical structure constructed in these type of words makes one phenomenon apparent. Examples containing no GS position display a sequence of three adjacent strong nodes, whereas examples where a GS position is present display no such sequence. We now concentrate on the cases where no GS is present. The adjacency of elements belonging to the same category has long been known as the perfect environment for the application of the Obligatory Contour Principle (Leben 1973). In metrical terms, the resolution of stress clash calls for a contour in the clashing string whereby the weakening of one of the nodes takes place, although the actual triggering environment for the weakening varies across languages and the notation used for this process also varies according to linguistic frameworks. Liberman & Prince (1977) call it a Rhythm Rule and Selkirk (1984) calls it Beat Movement.

According to empirical observations across languages (Hogg (1987) p. 148, and Halle & Vergnaud (1987)), the eurhythmic pattern where periodicity, or equal intervals, are present seems to be the most preferable in languages. This seems to be the case in BP as well. We propose that BP has a Rhythm Rule[58] such as the following:

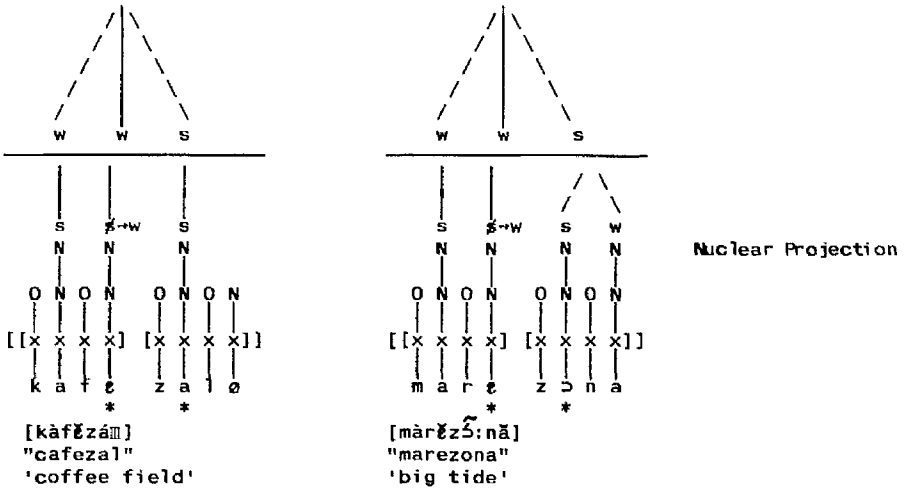
(197) Rhythm Rule in BP.

s s s → s w s

Although different forms of rhythm rules have been proposed and the actual way in which clashes are resolved in each linguistic framework is also different (i.e. grid theory, arboreal theory, etc.), the context and the solution in (197) apply to languages like English and to some types of clashes in Italian (Nespor & Vogel, 1989).

The application of the Rhythm Rule in (197) to the derivation of "cafezal" and "marezona" will result in the strings below:

(198)



With the application of the rhythm rule in (197), the word tree is constructed on the adjusted nodes.

[58]The term "Rule" as used here refers to a process that, like all others proposed, only occurs when it is motivated by the context.

The metrical analysis just proposed predicts non-primary stress on -ka- and on -ma- in "cafezal" and in "marezinha", respectively. It also predicts no stress on -fɛ- or on -rɛ-, and primary stress on -za- and on -zɔ-, respectively, by the application of the rhythm rule in (197). This corresponds to the actual realisation of these words, as shown below:

(199)	[kàfɛ̆zá]	"cafezal"	'coffee fields'
	[màrɛ̆zõ:nă]	"marezona"	'big tide'

We have now arrived at an explanation for the stress of both forms where the initial stem has two nuclei and for any other types of stem. We have seen that by analysing the suffixes "-zal", "-zona", "-zense", and "-zeiro" as analytic we can predict the correct pattern of stress for derivations where they are involved.

In this chapter we have analysed nouns and adjectives whose internal morphological structure is relevant from a phonological point of view. In these structures, we have seen that processes apply whenever their structural environment is met and that this happens in an independent fashion, i.e. the assignment of stress and the application of the rhythm rule. We have seen evidence that stress respects the Principle of Strict Cyclicity in all cases, and that whenever a clash is created it is resolved by the application of a rhythmic process whereby an alternating pattern of strong and weak nodes (s w) is re-established.

8.1 *Introduction.*

In this chapter we will attempt to prove that there are two distinct diminutive suffixes in BP, which surface [iñ] and [ziñ]. We will argue that the -iñ- form is non-analytic whereas the -ziñ- one is analytic. We will base our argument on phonological as well as morphological aspects of the two forms. The data used in this analysis come from the dialect of Natal. This somewhat special treatment given to the diminutive is motivated by the claim (Da Silva 1988) that there is only one diminutive suffix which surfaces in two different ways, [iñ] and [ziñ]. Due to Da Silva's claim and its considerable productivity in present-day Portuguese spoken in Brazil, we will analyse the diminutive based on data from Natal.

8.2 *Some Issues Related to the Diminutive.*

As mentioned above, there are two surface forms for the diminutive in BP, [iñu/a] and [ziñu/a]. Examples with these two forms are given below:

(200) a. Diminutives in -iñu/a-.

[kàzĩ:ñă]	"casinha"	'little house'
[bisiklɛ̃tĩ:ñă]	"bicicletinha"	'little bicycle"
[vɛ̀tɐ̃rănĩ:ñŭ]	"veteraninho"	'little veteran"
[kàrɔ̃nĩ:ñă]	"caroninha"	'little lift'
[kɐ̃mĩ:ñă]	"caminha"	'little bed"

b. Diminutives in -ziñu/a-.

[kàfɛ̃zĩ:ñŭ]	"cafezinho"	'little coffee'
[màràkŭžàzĩ:ñŭ]	"maracujazinho"	'little maracujá (fruit)'
[àbàkăšizĩ:ñŭ]	"abacaxizinho"	'little pineapple'
[sìpɔ̃zĩ:ñŭ]	"cipozinho"	'little stick'
[màrɛ̃zĩ:ñă]	"marezinha"	'little tide'

We start the analysis of these forms by arguing that the -iñ- form of the diminutive is non-analytic and that the -ziñ- one is analytic. Let us check whether this analysis is a plausible one in metrical terms.

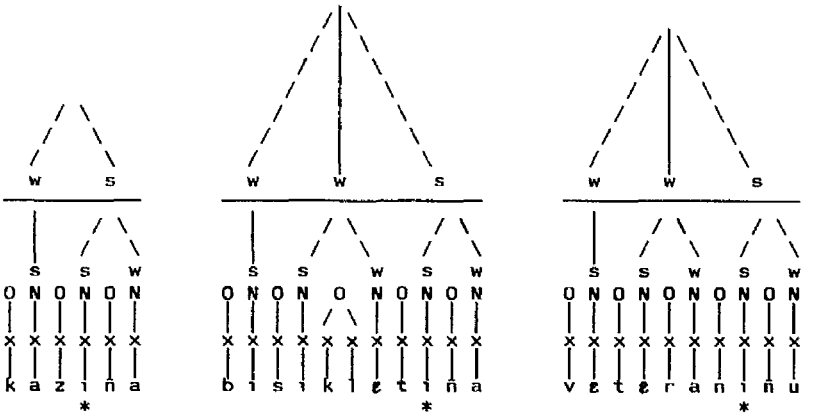
Following our proposals concerning the derivation of nouns and adjectives in BP in chapters 6 and 7, it follows that, once we

argue that a certain suffix is non-analytic, it should be attached to the stem or derivational stem before a GS can be added. If that is the case, it also follows that stress should be assigned to the whole string as if no morphological boundary existed. We are, therefore, considering the case of -iñ-. Before we can proceed to show how stress is assigned to these forms, let us remind ourselves of the metrical settings proposed for BP, which are given below:

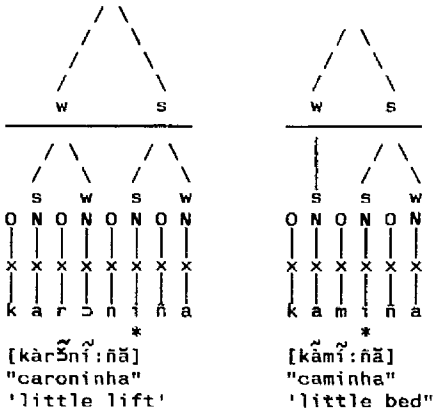
(201) i. Feet are binary and left-headed.
ii. Foot construction begins at the right edge of the word.
iii. Word trees are right-headed.
iv. BP is quantity insensitive.

Let us now apply these settings to the words in ((200)a) and check whether the metrical structure constructed in this fashion predicts the attested pattern for these words[59].

(202)



[59]In the following metrical structures, unless we are dealing with analytic derivations, only the final lexical marking is given. In analytic derivations, both lexical markings and their lexical gender information are given.



As shown above, the construction of metrical structure to the words in ((200)a) in a non-analytic fashion predicts the correct pattern of stress.

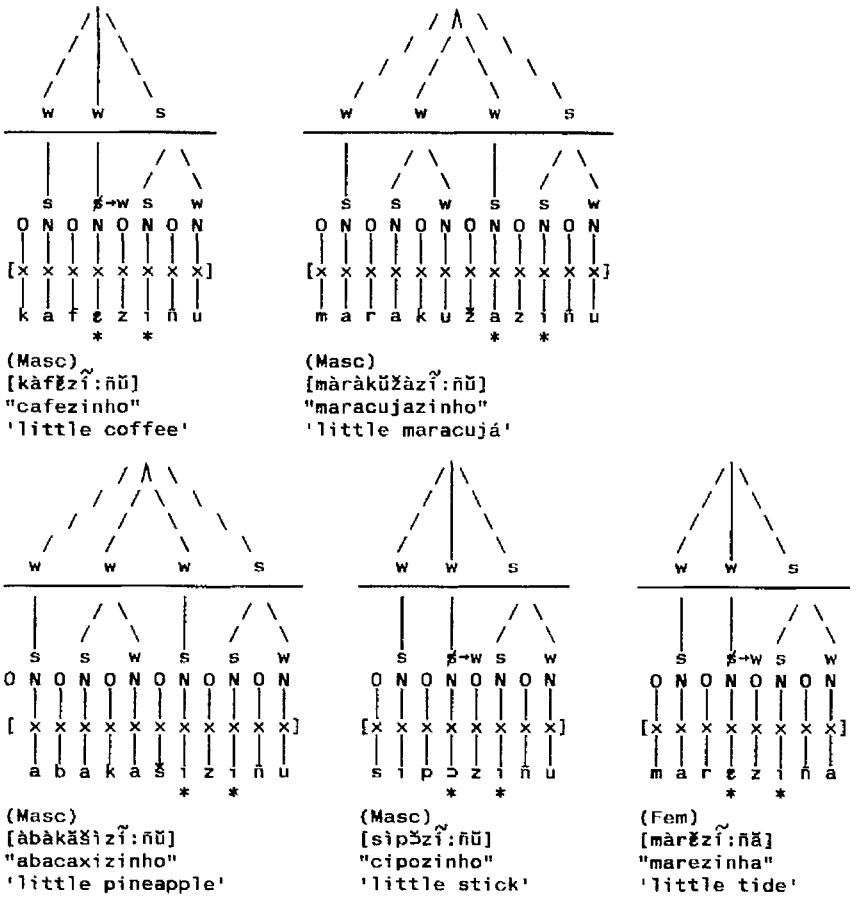
Since we propose that the -ziñ- form of the diminutive is analytic, it follows that it should be affixed to another domain and that the GS position should remain, if a GS is present (in parentheses in (203)). These forms would display the following type of domain organisation:

(203) [[stem+(GS)] [-ziñu/a]]

We have seen that, in this type of structure, stress is assigned to each domain in an independent fashion. After that, the domain boundaries of each individual domain are deleted and the larger domain is reached. The word tree is then constructed.

Let us now apply the metrical settings in (201) to the words in ((200)b) in an analytic fashion and check whether this predicts the attested pattern of stress. The structures below represent the stage of the derivation where stress has been assigned to the two individual domains and their individual boundaries have been deleted.

(204)



As shown above, the metrical structure constructed in an analytic fashion predicts the attested pattern of stress of the words in ((200)b). As far as the words whose stem consists of two nuclei are concerned ("cafezinho", "cipozinho", and "marezinha"), we notice that the rhythm rule proposed in chapter 7 applies in the environment where three strong nodes are found in a sequence. The application of the rhythm rule results in the weakening of the medial node. The word tree is then constructed.

The metrical structures in (202) and (204) prove that, with respect to stress, the interpretation of -iñ- and -ziñ- as non-analytic and analytic, respectively, is not problematic.

We shall now shed light on other aspects of the diminutive which we believe could support our argument.

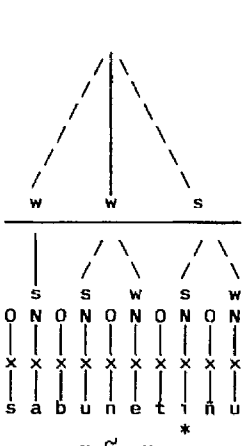
We start with the semantic aspects of these two forms. Before we can examine this aspect in diminutival expressions, it is important to note that certain words admit either of the two forms of the diminutive. This is the case in words whose GS is overt. Although the actual reason why this should be so is only discussed in section 8.3, here we will consider the most striking difference between the two forms. If we intend to propose that the two forms derive from distinct lexical items, we should expect some sort of semantic distinction between them. In that respect, it is interesting to check the meaning of the following forms:

(205)	[sàbŭné:tĭ]	[sàbùnĕtĭ̃:ñŭ]	'little soap'
		[sàbŭnètĭzĭ̃:ñŭ]	'bad soap'
	[kàtărá:tă]	[kàtàrătĭ̃:ñă]	'little waterfall (affectionate)'
		[kàtăràtăzĭ̃:ñă]	'little waterfall (depreciative)'

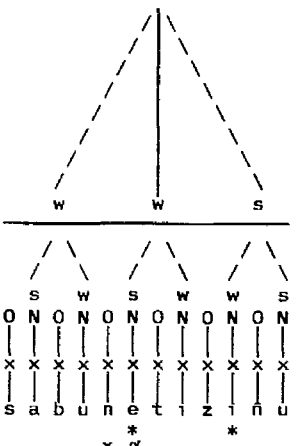
As shown above, the cases where -iñ- is derived always results in the diminutive of the initial semantic idea of the stem or, at most, in an expression of affection for the initial idea. On the other hand, the cases where -ziñ- is derived show that a pejorative meaning is given to the initial semantic idea of the entry.

The semantic aspect of the diminutival expressions above indeed suggests that we are dealing with two distinct lexical items. If this were not the case, we should only be able to derive one semantic idea. We must now check that this is interpretation is confirmed by metrical structure. The metrical structure of the forms in (205) is given below, where again in the analytic derivations (-ziñ-) the larger domain has already been reached:

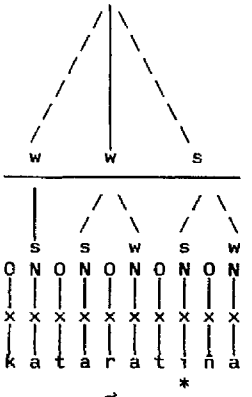
(206)



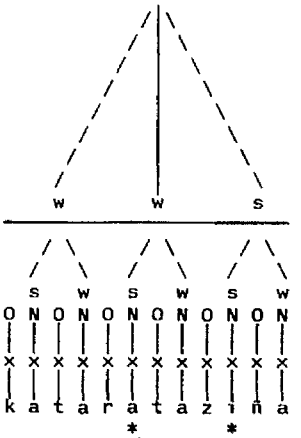
[sàbùnětĩ:ñŭ]
"sabonetinho"
'little soap'



[sàbŭnètĭzĩ:ñŭ]
"sabonetezinho"
'bad soap'



[kàtàrătĩ:ñă]
"cataratinha"
'light waterfall (affect.)'



[kàtăràtăzĩ:ñă]
"cataratazinha"
'little waterfall (deprec.)'

As illustrated above, the interpretation of -iñ- and -ziñ- as two independent lexical entries is supported by metrical structure - i.e. the resulting stress patterns are those predicted by the metrical analysis proposed here.

There are other aspects of the "diminutive" which can support the analysis that we are dealing with two lexical entries which are both called the diminutive. For example, we can examine the behaviour of each form of the diminutive in relation to certain morphological aspects of lexical items. In particular, let us consider the GS content of lexical items when each form of the diminutive is added. Since we propose that -iñ- and -ziñ- represent two distinct lexical entries, it should follow that one of

them might treat the GS spell-out of the lexical entry differently from the other. Examples of all possible types of derivation are given below.

(207)	LEXICAL ENTRY	-iñ-	-ziñ-
a.	[mìnĩ:nŭ]	[mìnĭnĩ:ñŭ]	[mìnĩnŭzĩ:ñŭ]
	*	*	* *
	"menino"	"menininho"	"meninozinho"
	'boy'	'little boy'	'any boy found in the streets'
b.	[kàbé:să]	[kàbĕsĩ:ñă]	[kàbèsăzĩ:ñă]
	*	*	* *
	"cabeça"	"cabecinha"	"cabeçazinha"
	'head'	'little head'	'stupid person'
c.	[bàrá:ʎŭ]	[bàrăʎĩ:ñŭ]	[bàràʎŭzĩ:ñŭ]
	*	*	* *
	"baralho"	"baralhinho"	"baralhozinho"
	'deck (cards)'	'little deck'	'poor quality deck'
d.	[sìdá:dĭ]	[sìdădĩ:ñă]	[sìdàdĭzĩ:ñă]
	*	*	* *
	(Fem)		
	"cidade"	"cidadinha"	"cidadezinha"
	'city'	'small town'	'bad small town'
e.	[lègũ:mĭ]	[lègŭmĩ:ñŭ]	[lègũmĭzĩ:ñŭ]
	*	*	* *
	(Masc)		
	"legume"	"leguminho"	"legumezinho"
	'vegetable'	'nice little vegetable'	'bad vegetable'
f.	[žìrăsɔ̀ɯ]	[60]	[žìrăsɔ̀ɯzĩ:ñŭ]
	*		* *
	(Masc)		
	"girassol"		'little vegetable' or
	'sunflower'		'bad vegetable'
g.	[àhtístă]	[àhtĭstĩ:ñă]	[àhtístăzĩ:ñŭ] (Masc)
	*	*	* *
	(FC)	(FC)	[àhtistăzĩ:ñă] (Fem)
			* *
	"artista"	"artistinha"	"artistazinho/a"
	'artist'	'a certain artist'	'bad artist'
h.	[kàbé:să]	[kàbĕsĩ:ñă]	[kàbèsăzĩ:ñŭ] (Masc - H)
	*	*	* *
	(FC)	(FC)	[kàbèsăzĩ:ñă] (Fem - H)
	"cabeça"	"cabecinha"	"cabeçazinho/a"
	'mastermind (plot)'	'little mastermind'	'stupid mastermind'

(207) illustrates all possible combinations of spell-out for the GS position. ((207)a) represents cases where the speaker can choose the gender and applies the basic rule of the spell-out of -a for the feminine and -u for the masculine (FC). In this case, we can see that the same information appears at the end of the only domain, in the derivation of -iñ-, and twice in the derivation of

[60]The derivation with -iñ- is unacceptable for BP speakers.

-ziñ-. The next two examples (((207)b)-((207)c)) are cases where the gender of the derivative does not follow the basic rule. The masculine/feminine distinction is not possible. In this type of lexical entry, the speaker knows that he/she should omit the GS spell-out of the lexical entry, add -iñ-, and only then should he/she allow the GS spell-out to be realised, i.e. at the end of the only domain of the derivation. When -ziñ- is the chosen form, the speaker knows that he/she should leave the lexical entry as it is in the lexicon, and then add the corresponding default morphology gender information at the end of the final domain. The examples in ((207)d), ((207)e) and ((207)f) represent cases where the GS spell-out does not follow default morphology, i.e. the GS position is filled with -i, -i, and, in the case of ((207)f), no GS is present. In these derivations, default morphology applies to all GS positions according to the gender information of the original lexical entry.

So far, the two forms of the diminutive have been seen to behave in exactly the same way. The GS position which follows either of the forms receives the same content, regardless of what lexical information is given. We now consider the cases where [a] occupies the GS position (((207)g)-((207)h)). In the derivation of -iñ-, -a fills the GS position. As for the derivation of -ziñ-, -a fills the GS position in the leftmost domain, whereas the final GS position follows default morphology. It is interesting to note that in the derivation of -iñ-, the GS position is filled with -a regardless of the gender chosen by the speaker. The derivations in (((207)g)-((207)h)) prove that the two forms of the diminutive treat the -a type of GS spell-out differently. This fact provides us with the third piece of evidence for the different behaviour of -iñ- and -ziñ-.

So far we have seen three types of evidence in support of the interpretation of -iñ- and -ziñ- as two lexically distinct suffixes. The first type of evidence is the fact that the correct pattern of stress is predicted if -iñ- is interpreted as non-analytic and -ziñ- is interpreted as analytic. The second type of evidence stems from the fact that the derivation of -iñ- and of -ziñ- from a

single suffix would predict that both forms should mean exactly the same, which is not the case. The third type of evidence is related to the difference in the treatment that -iñ- and -ziñ- give to the content of the GS position of the initial lexical entry. Whilst -iñ- allows for the -a type of spell-out of the GS to be realised, the -ziñ- suffix only allows the default morphology type of spell-out to fill in the GS which follows it.

One further piece of evidence can be obtained as we analyse the derivation of a special meaning that can derived by doubling the so-called "diminutive". By doubling the diminutive suffix, it is possible to express extremely exaggerated ideas such as "extremely small" or "extremely bad". Let us consider the following data:

(208) a. [kàhĩñũzĩ:ñũ] "carrinhozinho" 'extremely bad car'

b. * [kàhũzĩñũzĩ:ñũ]

c. * [kàhũzĩñĩ:ñũ]

d. * [kàhĩñĩ:ñũ]

e. [kàfɛzĩ:ñũ] "cafezinho" 'little coffee'

f. * [kàfɛĩñũzĩ:ñũ]

g. * [kàfɛzĩñũzĩ:ñũ]

Before we can consider the relevance of the data in (208), it is important to learn that the diminutive with -iñ- is not possible in words which have no GS[61], as in the case of "café" 'coffee'. This rules out the derivation in ((208)f). It is also important to be aware that, for reasons which do not follow from this analysis, the sequence 'palatal consonant + palatal vowel' is highly constrained in BP, especially word-finally. Although the precise circumstances under which this sequence is not permitted seem unclear, here we will simply point out that [ñi] is not allowed if -i- is primarily stressed. This constraint applies more rigidly when two or more of these sequences appear in succession. This means that ((208)c) and ((208)d) are excluded. Not only do we now need to explain

[61]But this is only discussed in section 8.3.

how the double diminutive is assigned, but we must also explain why ((208)b) and ((208)g) are excluded.

Let us suppose that these extremely exaggerated ideas are derived by adding -iñ- and then -ziñ- to a lexical entry, in this strict order. If the order in which these forms are added is a fixed one, it follows that whenever the concatenation of the first form is impossible with a specific entry, the double diminutive is likewise impossible. This hypothesis seems to explain why ((208)f) is excluded and why only one form of the diminutive is possible in certain derivations, as shown in ((208)e). Also, if we limit the double diminutive to the application of -iñ- and -ziñ- in this strict order, it follows that concatenations where either -iñ- or -ziñ- are derived twice are excluded. This explains why ((208)b), ((208)d) and ((208)g) are excluded. From this it follows that the reverse order of the suffixes is excluded ((208)c). To sum up, the assertion that the double diminutive is derived through the concatenation of -iñ- and -ziñ- in this strict order predicts that the only case where this extremely exaggerated idea is possible is in entries where -iñ- is compatible.

This analysis predicts that stems that admit -ziñ- as the only possible diminutive can never be given the exaggerated meaning of "extremely small" or "extremely bad". This prediction turns out to be true. In this analysis, all we need to say is that the idea of 'exaggerated smallness' or 'extremely low quality" is given by the application of the non-analytic suffix followed by the analytic one in this strict order. This automatically explains why the so-called 'doubling' of the diminutive is only possible in words where -iñ- is possible, and why it has to be excluded from words where only -ziñ- is possible. A stem that can take -iñ- will necessarily take the sequence -iñ+GS+ziñ+GS- as well. As for the cases where only -ziñ- is possible, they will not have a double diminutive form. The so-called 'doubling' of the diminutive constitutes the fourth motivation for the interpretation of -iñ- and -ziñ- as two distinct suffixes.

In the next section we will attempt to account for the assignment of the diminutive in the dialect of Natal.

8.3 *The Diminutive in Natal.*

We have given several reasons why -iñ- and -ziñ- must be interpreted as deriving from two distinct lexical entries based on data from Natal. In this section we will attempt to show how each one of these forms is assigned in Natal.

The assignment of the diminutive takes place according to the type of lexical item at issue. For example, if the GS follows the basic morphological rule of -a for the feminine and -u for the masculine, it will accept the -iñ- version of the diminutive, as shown below:

(209)	[bãnã́:nă]	[bãnãnĩ́:ñă]
	*	*
	"banana	"bananinha"
	'banana'	'little banana'
	[mɔ̃rɛ̃́:nŭ]	[mɔ̃rɛ̃nĩ́:ñŭ]
	*	*
	(FC)	(FC)
	"moreno"	"moreninho"
	'brunette'	'nicely sun-tanned'

As has been shown throughout this chapter, nouns and adjectives which follow the basic morphological rule for the spell-out of the GS position will invariably accept the -iñ- version of the diminutive.

Let us now reconsider the cases where the reverse of the basic morphological rule applies, namely those cases where -u is the GS spell-out of a feminine word and -a is the GS spell-out of a masculine word. These cases are illustrated below:

(210)	[trí:bŭ]	[trìbĩ́:ñă]
	*	*
	"tribo"	"tribinha"
	'tribe'	'little tribe'

```
[kàbé:să]              [kàbĕsĩ́:ñă]
   *                       *
"cabeça"               "cabecinha"
'mastermind (plot)'    'little mastermind'
```

A word needs to be said about the derivation of [trìbíñă]. This seems to be the only word in Portuguese whose gender is feminine and whose GS spell-out is [u]. The derivation of the diminutive in -iñ- out of this stem makes speakers frequently think and report that neither the form where [u] fills the GS position nor the one where [a] occupies this position sounds correct. Although they feel reluctant to say [trìbíñŭ], they also say that [trìbíñă] is unacceptable. In this conflicting situation, speakers usually find it preferable to use the -ziñ- form of the diminutive whereby a more readily acceptable form is achieved, i.e. [trìbŭzíñă], where [u] fills the first GS position and [a] fills the last one. As far as the reverse situation is concerned, the lexical spell-out fills the GS position in the derivation of -iñ-.

The final case where -iñ- is compatible is that in which where -i fills the GS position. These cases are illustrated below (repeated from (207):

```
(211)  [sìdá:dĭ]          [sìdādĩ́:ñă]
           *                    *
       (Fem)
       "cidade"           "cidadinha"
       'city'             'small town'

       [lègṹ:mĭ]          [lègŭmĩ́:ñŭ]
           *                    *
       (Masc)
       "legume"           "leguminho"
       'vegetable'        'nice little vegetable'
```

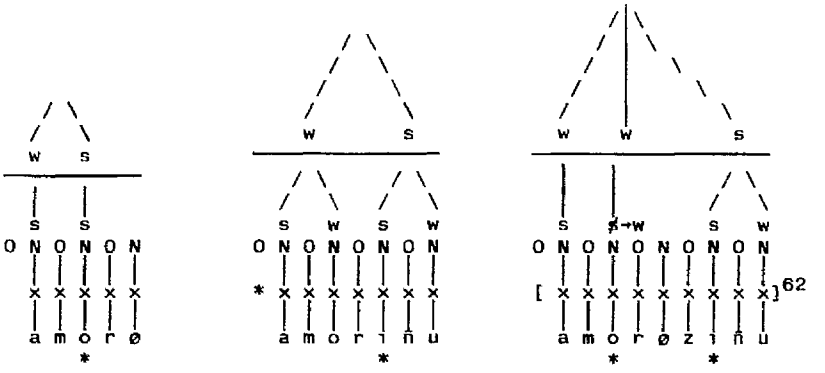
The cases where -i fills the GS position also show that the -iñ- version of the diminutive is compatible.

Let us now consider the phonological similarities between these cases. Phonologically, what these cases have in common is the fact that the stem in all cases ends in an onset and that the GS position is filled. The derivation with -iñ- is, however, incompatible in those cases where the GS position is absent, regardless of whether the stem ends in a nucleus or in a

consonant. In these cases, the only possible diminutive is -ziñ-. We can also add that -ziñ- is always possible in all types of lexical item, as long as the integrity of this lexical item is maintained - that is, the entry's GS spell-out is maintained as listed in the lexicon. These facts seem to suggest that there is a relation between the absence of a GS position and the possibility of combination of -iñ- or -ziñ-.

These facts about the derivation of the diminutive seem to suggest that a phonological condition in these cases does not hold, otherwise all of the following derivations should be possible.

(212)



As shown above, words whose stem ends in a consonant but have no GS are incompatible with -iñ- even though phonologically nothing would prevent their combination. The sequence consisting of a stem final onset + suffix initial nucleus would undoubtedly satisfy the Licensing principle. Since these derivations are unacceptable in all dialects of BP[63], we are forced to conclude that speakers learn that lexical items with no GS are incompatible

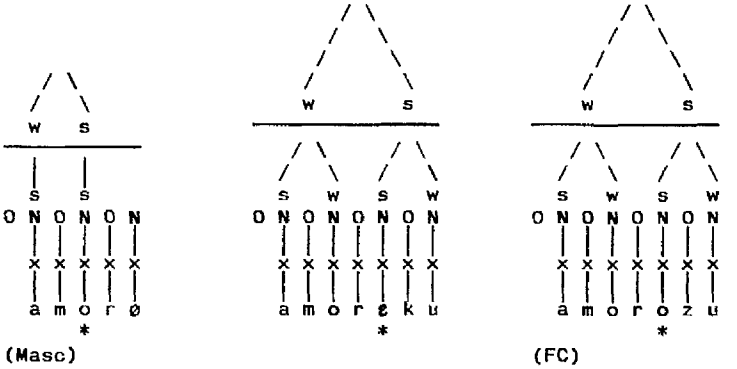
[62]As in the preceding cases, this structure illustrates the stage where feet have already been constructed on each individual domain and their (individual) boundaries have been deleted.

[63]Many derivations of this type are acceptable in European Portuguese.

with the diminutive -iñ-. In fact, non-natives usually produce morphological errors whilst attempting to use the diminutive in Portuguese. The diminutive of these lexical items is only achieved by adding -ziñ-, which is analytic and hence independent of the preceding domain.

This position, however hard to accept, explains why other non-analytic suffixes are allowed with lexical items which have no GS but whose stem ends in a consonant. This restriction is limited to the diminutive. Derivations with other suffixes where no GS is present in onset final stems are illustrated below:

(213)



The diminutive in Natal will depend, as suggested by the data, entirely on whether the GS is present or not.

For the reasons discussed above, we would like to propose that the two surface forms -iñ- and -ziñ- be lexically represented as two distinct lexical entries, the first of which (-iñ-) is non-analytic, and the second one (-ziñ-) analytic.

Let us now concentrate on the -ziñ- form of the diminutive. As mentioned at the beginning of this chapter, -ziñ- can also be used with the pejorative connotation. If that is the case, it means that all stems will accept this pejorative suffix in an analytic fashion. Since some stems accept -ziñ- as the only possible diminutive, there is a strong prediction that we can make. Specifically, we

can predict that whenever only -ziñ- is possible, it is semantically ambiguous. This is indeed true, as shown below:

(214)	[nàtãɯzĩ:ñũ]	"Natalzinho"	'little Christmas' or 'bad Christmas'
	[kãmărà̀ɯzĩ:ñũ]	"camarãozinho"	'little shrimp' or 'bad shrimp'

We now discuss the actual acceptability of the -ziñ- diminutive in Natal. Because in principle -ziñ-, as an analytic suffix, is phonologically compatible with any stem, we need to find some way of excluding -ziñ- as the diminutive in Natal in cases where the stem accepts -iñ- as a possible diminutive. A simple way to exclude these possible results is to incorporate the following statement into the derivation of the diminutive in Natal only.

(215) ***The Derivation of the Diminutive in Natal.***

Add the analytic form of the diminutive (-ziñ-) if and only if the non-analytic (-iñ-) is not possible.

This means that those derivations which carry the meaning of smallness, lightness or affection in (216) below are excluded.

(216)	* [kàhũzĩ:ñũ]	'small little car'
	* [màlètăzĩ:ñă]	'small little suitcase'

The selection of which version of the diminutive is suitable for a given stem depends entirely on whether the GS is present in a lexical entry or not, even if its stem ends in a consonant.

As shown above, the assignment of -ziñ- is unacceptable in Natal if the entry can take -iñ-, and -iñ- is unacceptable if there is no GS.

In this chapter, we have discussed the derivation of the diminutive in Natal. We have proposed an analysis based on the Natal data where the diminutive is lexically represented by two lexical entries, -iñ- and -ziñ-, which are non-analytic and analytic, respectively.

9.1 *Introduction.*

As has been mentioned throughout this thesis, in this chapter we will not present a full analysis of the surface antepenultimate stress pattern in nouns and verbs. This discussion summarises the factors that seem to play a role in the assignment of antepenultimate stress. A full analysis is to be carried out in a subsequent project.

9.2 *Some Facts About Antepenultimate Stress.*

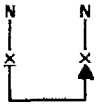
The discussion that follows draws from some facts about Italian, subsumed under the label "open syllable lengthening" in traditional Italian grammars (J. Kaye, PC). In Italian, primary stressed open syllables obligatorily lengthen in penultimately stressed words. Primary stressed nuclei in antepenultimately stressed words, on the other hand, are not allowed to lengthen. This is also true for BP, as shown below:

(217)	a.	[mé:za] *		"mesa"	'table'
		[está:du] *		"estado"	'state'
		[bɔ́:la] *		"bola"	'ball'
	b.	[tátika]	* [tá:tika]	"tática"	'tactics'
		[étika]	* [é:tika]	"ética"	'ethics'
		[álibi]	* [á:libi]	"álibi"	'alibi'

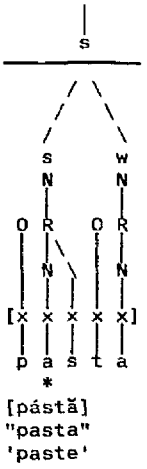
The examples in (217) represent a typical realisation in Natal. They show that an antepenultimate open syllable which receives primary stress is not allowed to lengthen. The similarity between these facts and those from Italian has been attributed to the role of metrical heads. It has been proposed that the head of a metrical domain intrinsically needs to govern. It must be an independent governor. (218) illustrates the type of constituent that is expected to occupy the head of the domain.

(218) a. 

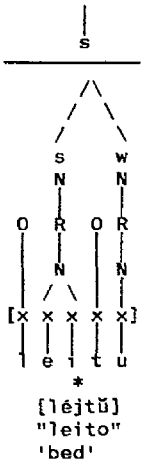
b. 

c. 

According to this analysis, the primary stressed nucleus must be involved in one of the governing relations illustrated in (218). It needs to govern locally, independently of the governing relation it establishes with the final nucleus of the word in the formation of a foot. This may be seen in the cases where the GS is required. In penultimately stressed words, if the primary stressed nucleus does not match any of the structures in (218), it lengthens. Let us illustrate this with the following:

(219) a. 

[pástă]
"pasta"
'paste'

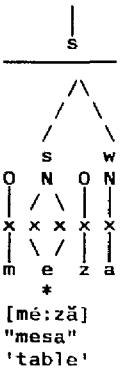
b. 

[léjtŭ]
"leito"
'bed'

Nuclear Projection

In (219), the head of the metrical domain is branching. Being a branching constituent, it satisfies the governing conditions imposed on a metrical domain head and therefore remains as it is. In lexically non-branching nuclei, on the other hand, this governing condition needs to be satisfied. The primary stressed nucleus needs to be an independent governor; therefore, it lengthens. (220) illustrates a case where a primary stressed nucleus is lengthened.

(220)



[mé:zǎ]
"mesa"
'table'

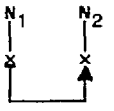
In antepenultimately stressed words, on the other hand, there is no need for the primary stressed nucleus to branch, since a nucleus is available between the primarily stressed nucleus and the weak node of that foot. Because the head does not branch (owing to the presence of the medial nucleus) the metrical head is said to have contracted a local governing relationship with the penultimate nucleus and satisfied its needs in terms of government. To account for similar facts in the Natal data, we adopt this analysis and propose the following hypothesis:

(221) N_1 N_2 N_3

'Let N_1 be the primary stressed nucleus in an antepenultimately stressed word and N_2 and N_3 the two posttonic nuclei, there is a governing relationship between the primary stressed nucleus (N_1) and a posttonic N_2 non-branching nucleus. Metrical heads need to govern.

With this hypothesis, we are proposing that the primary stressed nucleus (N_1) of an antepenultimately stressed word metrically governs not only the final nucleus (N_3), as a weak member of the foot, but also the penultimate one (N_2) in a more local fashion. The governing relation holding between N_1 and N_2 will have the structure in ((218)c) (repeated below).

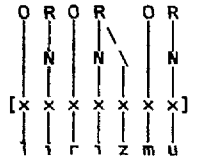
(222) a.



This means that the segment under N_2 must satisfy the conditions of a governee. This constituent must be non-branching, for example. Antepenultimate stress would therefore never occur in

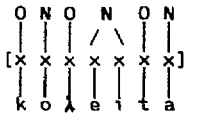
words of the following type:

(223) a.



[lìrízmŭ]
"lirismo"
'lyrism'

b.

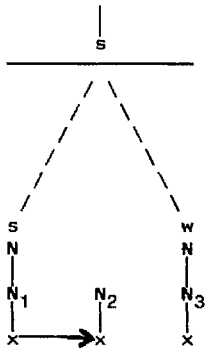


[kòʎéjtă]
"colheita"
'crop'

In (223), the penultimate nucleus could never be a governed segment because it branches.

An antepenultimately stressed word would display the following type of metrical structure:

(224)



Crucial to the structure above is the fact that in governing N_2, N_1 satisfies the governing requirements for domain heads. It follows that the non-branching nucleus under N_1 will not be allowed to lengthen.

We have proposed that the primary stressed nucleus governs the medial nucleus in antepenultimately stressed words. It is crucial now to give evidence for the relationship between N_1 and N_2, since we already know from our analysis that the metrical head of the word governs the final nucleus (N_3). The presentation of such evidence will be our main goal in this discussion.

Let us therefore consider the governing domain N_1-N_2. The first type of evidence we can provide is related to constituent structure. If N_1 indeed governs N_2, then it would be expected

that there should be some constraints imposed upon N_2 with respect to what type of constituent is permitted. For example, N_2 would not be expected to branch. We mentioned at the beginning of this section that the nucleus under N_2 is not allowed to branch in BP. This means that the following type of words are excluded in BP:

(225)	* taraika	but, [pérʃlă]	"pérola"	'pearl'
	* lirismu	but, [lirísmŭ]	"lirismo"	'lyrism'

The second type of evidence for the governing relationship between N_1 and N_2 has to do with which segments can occupy the N_2 position. Since N_1 bears primary stress (= governs), it follows that the complexity of N_2 should in some way be restricted by the complexity of N_1. We will now investigate the types of segments that can arise in N_2.

In fact, the Natal dialect can provide us with just the type of evidence we need. As an example, let us consider which segments are allowed in N_2 according to which segment occupies N_1[64].

		N_1	N_2	Examples		
(226)	a.	i	i	[fízĭkă]	"física"	'physics'
			ɛ	[bípɛ̆dĭ]	"bípede"	'biped'
			a	[sílăbă]	"sílaba"	'syllable'
			ɔ	[ɛ̆kivʃkŭ]	"equívoco"	'mistake'
			u	[fístŭlă]	"fístula"	'fistula'
	b.	e	e	[pésĕgŭ]	"pêssego"	'peach'
			a	[bébădă]	"bêbada"	'drunk (fem.)'
			o	[ézŏdŭ]	"êxodo"	'exodus'
	c.	ɛ	i	[sétĭkŭ]	"cético"	'sceptical'
			ɛ	[férɛ̆trŭ]	"féretro"	'funeral'
			a	[pétălă]	"pétala"	'petal'
			ɔ	[pérʃlă]	"pérola"	'pearl'
			u	[sékŭlŭ]	"século"	'century'
	d.	a	i	[álĭbĭ]	"álibi"	'alibi'
			ɛ	[álʒɛ̆bră]	"álgebra"	'algebra'
			a	[lábărŭ]	"lábaro"	'flag'
			ɔ	[párʃkŭ]	"páraco"	'parish priest'
			u	[fábŭlă]	"fábula"	'fable'

[64]The possibilities under N_2 in (226) refer specifically to attested cases, although according to Natal speakers -o- and -a- are possible in the N_2 position in ((226)f).

e.	ɔ	i	[àpɔ́lĭsĭ]	"apólice"	'policy'
		ɛ	[kɔ́lĕră]	"cólera"	'cholera'
		a	[àlkɔ́lătră]	"alcólatra"	'alcoholic'
		ɔ	[àstrɔ́lɔgŭ]	"astrólogo"	'astrologer'
		u	[kɔ́pŭlă]	"cópula"	's e x u a l intercourse'
f.	o	e	[fólĕgŭ]	"fôlego"	'breath'
g.	u	i	[públĭkŭ]	"público"	'public (masc)'
		ɛ	[útĕrŭ]	"útero"	'womb'
		a	[búfălŭ]	"búfalo"	'buffalo'
		ɔ	[flúɔr]	"flúor"	'fluorine'
		u	[músculŭ]	"músculo"	'muscle'

The number of elements allowed in N_2 depends on the number of elements in N_1. Vowel quality matters in the sense that the segment under N_2 cannot consist of more elements than the segment in N_1. As shown in ((226)d), A^t, on the other hand, can govern any segment independently of its complexity. If N_1 is occupied by a complex but lax nucleus (((226)c) and ((226)e))[65], the segment under N_2 can consist of up to two elements in its representation. N_2 can be complex, in which case it is lax (-ɛ- or -ɔ-), or it can be tense, in which case it is simplex (-i- or -u-), but altogether it should not exceed the limit of two elements. By the same token, where a simplex tense nucleus ([i] or [u]) occupies the N_1 position (((226)a) and ((226)g)), N_2 can be occupied either by a complex lax nucleus (-ɛ- or -ɔ-), or by a simplex tense one (-i- or -u-). A good piece of evidence for the relevance of the number of elements involved in the representation of N_1 is the fact that A^t can be governed by either a complex lax nucleus (((226)c) and ((226)e)) or a simplex tense one (((226)a) and ((226)g))[66].

As for the case where N_1 is complex and tense (((226)b) and ((226)f)), there are gaps in the distribution of segments under N_2. These gaps are most probably accidental, since the existing cases coincide exactly with those where the number of elements is the maximum allowed under that position. There should be no reason

[65]The case where both the I° and the A^t lines are active.

[66]Here, what matters is the fact that it can be governed.

why exactly the cases where segments consisting of a lower number of elements were excluded. We will consider these gaps as being irrelevant to our discussion.

(226) can provide another type of evidence in favour of the relationship N_1-N_2. We refer to the occurrence of I^{+} in the two positions. If N_1 and N_2 are occupied by mid vowels, either I^{+} is present in both or in none. According to our discussion, there should be no reason why a complex tense nucleus N_1 would not be followed by a complex lax N_2. Cases such as those are illustrated below:

(227) [pésĕgŭ] * [pésɛgŭ] 'peach'

The fact that tenseness is either present or absent from both N_1 and N_2 reinforces our proposal for a closer relationship between N_1 and N_2. The agreement in tenseness between N_1 and N_2 follows the vowel harmony analysis. A primary stressed -e- would always spread I^{+} onto the preceding lax mid vowel in verbs. Following the same analysis, an antepenultimately stressed lax mid vowel cannot be followed by a tense one. The only requirement concerning the segment that occupies the N_3 position is that both lines of representation are not allowed to be active at the same time in the segment that occupies this position. This requirement applies to all unstressed word final nucleus in BP.

Let us sum up our results to this point. In this section we have argued that there is a governing relationship between the segments occupying the N_1 and N_2 positions in antepenultimately stressed words. We have given evidence that the relationship between N_1 and N_2 is different from the relationship between a primary stressed nucleus and the word final one (N_3). The relation holding between N_1 and N_2 satisfies the governing requirement for a domain head. In the construction of metrical structure, only the projected nuclei are accessible. This means that metrical structure is not aware of the existence of two nuclear positions under the primary stress node when it is

constructed. Metrical structure settings apply as they do to any other lexical entry and an antepenultimate pattern of stress is derived. This explains why the quantity insensitivity which is operative in BP is not contradicted by the fact that the antepenultimate pattern of stress is limited to words where the final two syllables are light. The medial nucleus (N_2), which is governed by the primary stressed nucleus (N_1) (and thus not allowed to branch), is not visible at the level where stress is assigned.

However, once we propose this, one question that arises immediately refers to how metrical structure can ignore the two nuclei below that node. At the present stage of the analysis, we have to assume that this is a lexical property of certain lexical entries. If antepenultimate stress were not assigned according to some lexical information, there should be no reason why this pattern of stress was absent in verbs. There is not a single case of antepenultimate stress in verbs, except for the cases displaying analytic morphology, as in -amos. As shown in our analysis of verbs, they invariably follow the penultimate pattern of stress. Also, there should no reason why the primary stressed nucleus in antepenultimately stressed words never lengthens.

This concludes our discussion on antepenultimate stress in BP. In this chapter we gave evidence for a closer relationship between the primary stressed nucleus and the following non-final one (N_1-N_2). It follows the governing requirements for domain heads. In penultimately stressed words this need is satisfied by the lengthening of the primary stressed nucleus if it is not branching lexically. In antepenultimately stressed words no such lengthening occurs. The primary stressed nucleus contracts a governing relation with this extra (medial) nucleus. This justifies the absence of lengthening in antepenultimately stressed words and, at the same time, guarantees that the construction of metrical structure will result in an antepenultimate pattern of stress. We have given altogether three types of evidence in support of this. The first type of evidence is related to the complexity of the

penultimate nucleus; it is not allowed to branch. The second type of evidence concerns which segments can occupy the penultimate nucleus according to the quality of the primary stressed one. We proved that the number of elements in the primary stressed nucleus determines exactly how many elements can occupy the penultimate nucleus. The final piece of evidence for the N_1-N_2 domain relates to the presence/absence of $\pm^{\dagger}$ in N_1 and N_2. There we noticed that although $\pm^{\dagger}$ could be absent from N_2 when N_1 is complex and tense, no such cases existed. This reinforces our proposal that the relation between N_1-N_2 is stronger than the relation between a primary stressed nucleus and the final nucleus.

10.1 *Conclusion.*

This thesis has contributed to the understanding of the phonology of Brazilian Portuguese for the reasons summarized below.

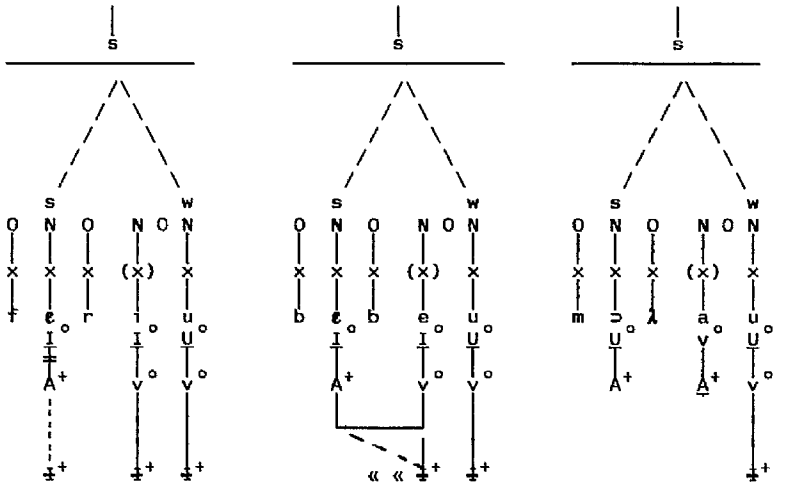
The assignment of stress to verbs, nouns, and adjectives in Brazilian Portuguese has been analysed in depth for the first time. The analysis has shown that stress in BP is not assigned lexically or by the application of rules, as proposed in earlier studies, but rather through a set of parameters available in the theory. It is important to note that this set of parameters is of a very simple nature. In this analysis a new level of metrical structure has been proposed, a level where rhythmic processes are observed.

The worth of this work remains to be verified when a full analysis of the remaining categories and an explanation for antepenultimate stress have been more successfully achieved.

A P P E N D I X

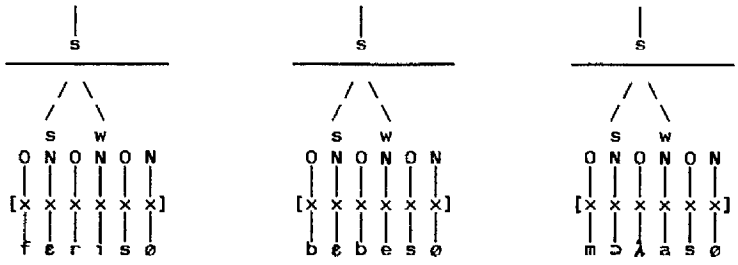
(1) The Indicative Present

P1



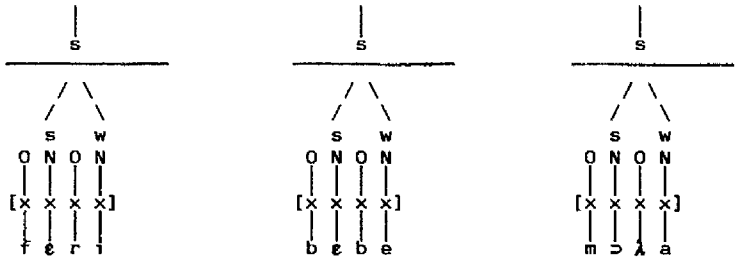
[fírŭ] [bébŭ] [mɔ́λŭ]

P2



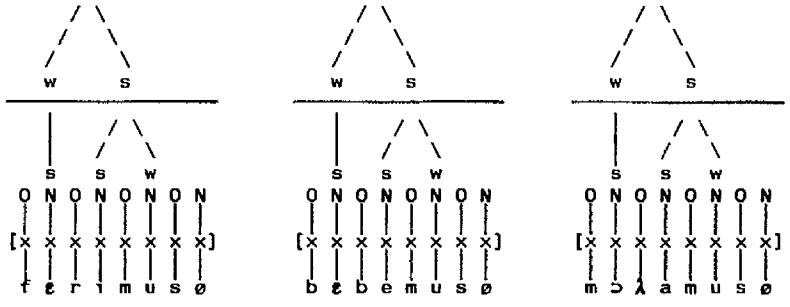
[férĭs] [bɛ́bĭs] [mɔ́λăs]

P3

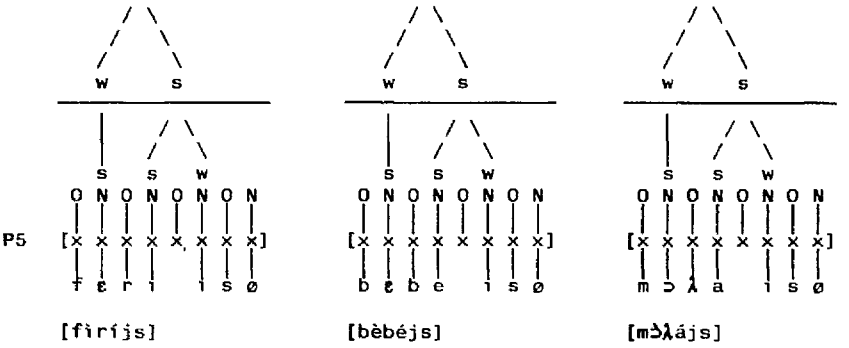


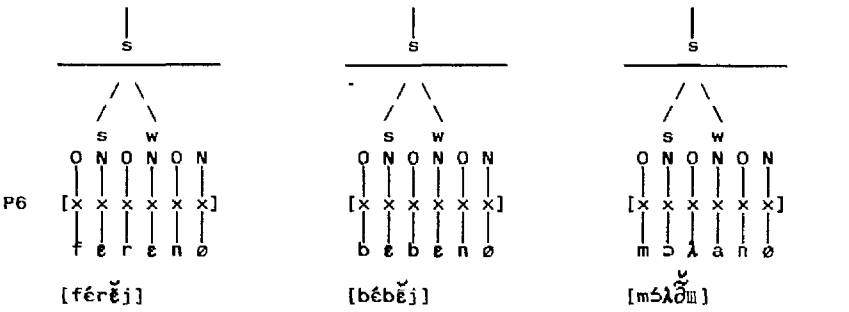
[fɛ́:rĭ] [bɛ́:bĭ] [mɔ́:λă]

P4

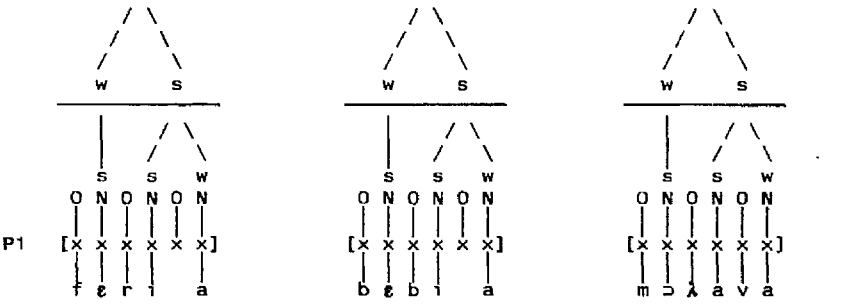


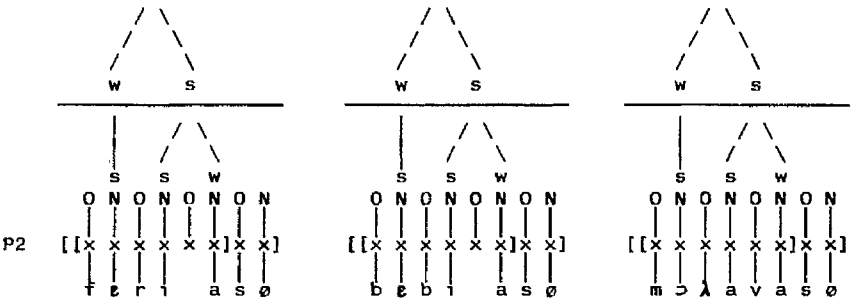
[firĩ́mũs] [bɛ̀bɛ̃́mũs] [mɔ̀λã́mũs]



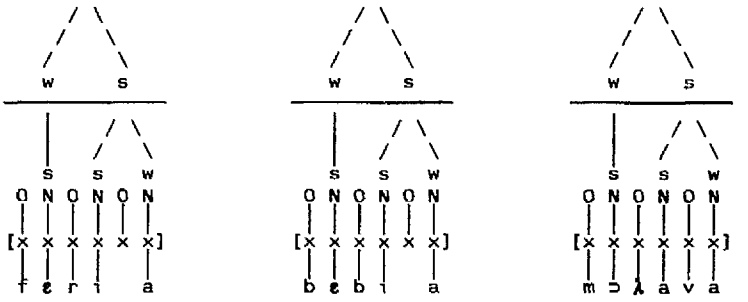


(2) The Indicative Imperfect.



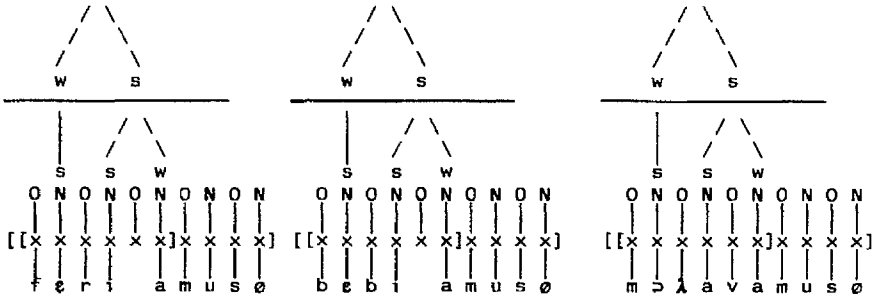


P3



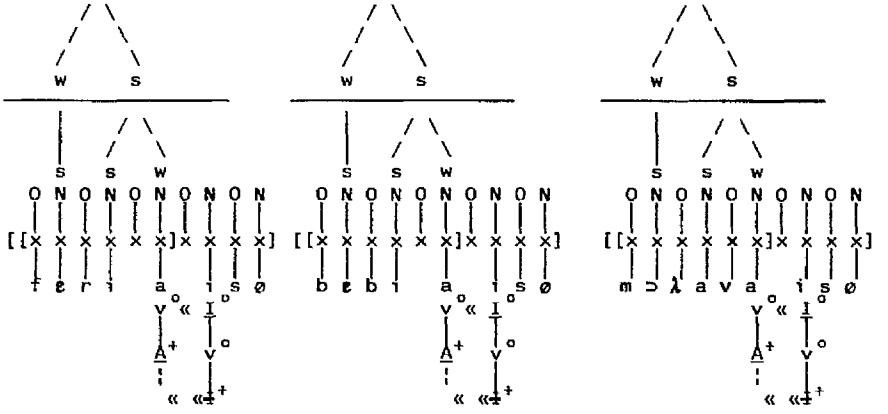
[fìríă] [bìbíă] [mɔ̀ɬávă]

P4



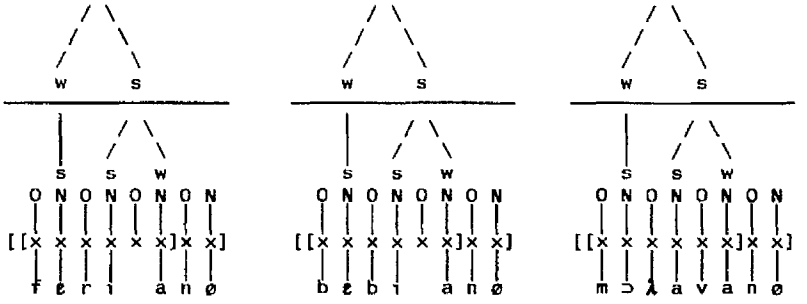
[fìríə̃mŭs] [bìbíə̃mŭs] [mɔ̀ɬávə̃mŭs]

P5



[fìríə̌js] [bìbíə̌js] [mɔ̀ɬávə̌js]

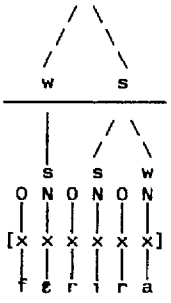
P6



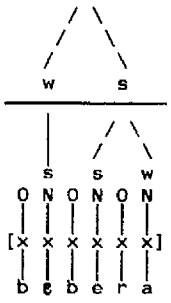
[fìríə̆ᵚ] [bìbíə̆ᵚ] [mɔ̀ɬávə̆ᵚ]

(3) The Pluperfect.

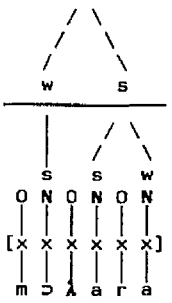
P1



[firírə̆]

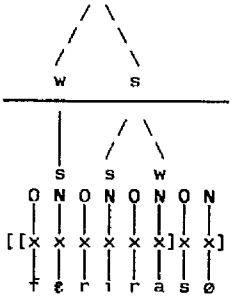


[bèbérə̆]

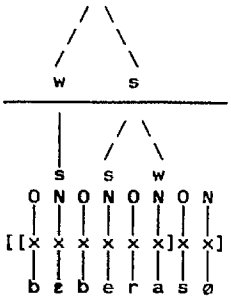


[mɔ̀ʎárə̆]

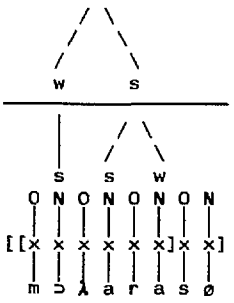
P2



[firírə̆s]

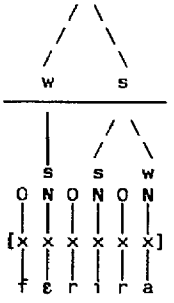


[bèbérə̆s]

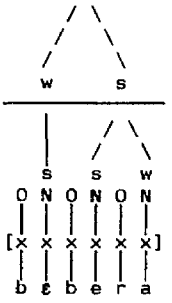


[mɔ̀ʎárə̆s]

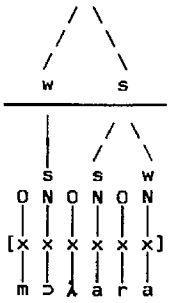
P3



[firírə̆]

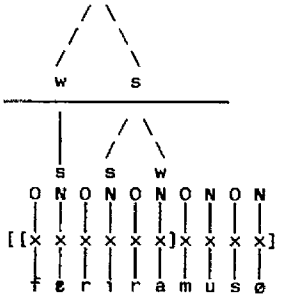


[bèbérə̆]

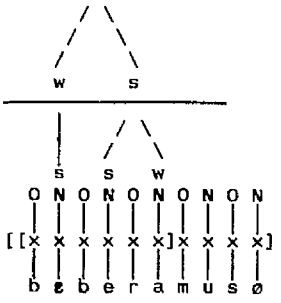


[mɔ̀ʎárə̆]

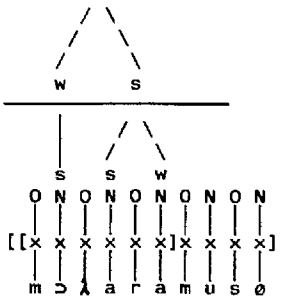
P4



[firírə̃mŭs]

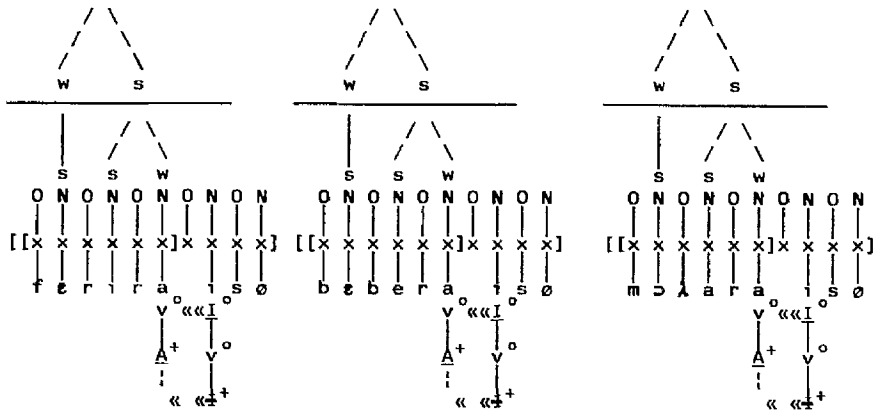


[bèbérə̃mŭs]



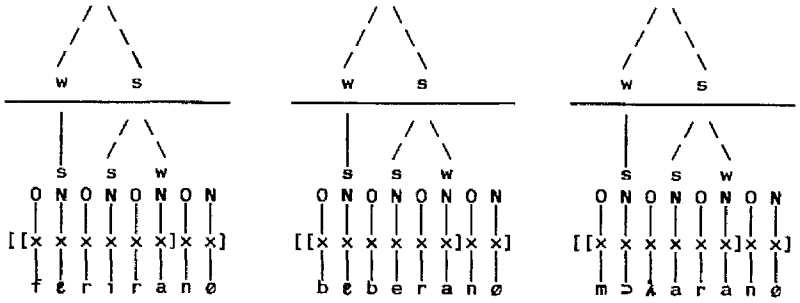
[mɔ̀ʎárə̃mŭs]

P5



[fìrírɛ̌ĭs] [bèbérɛ̌ĭs] [mɔ̀lárɛ̌ĭs]

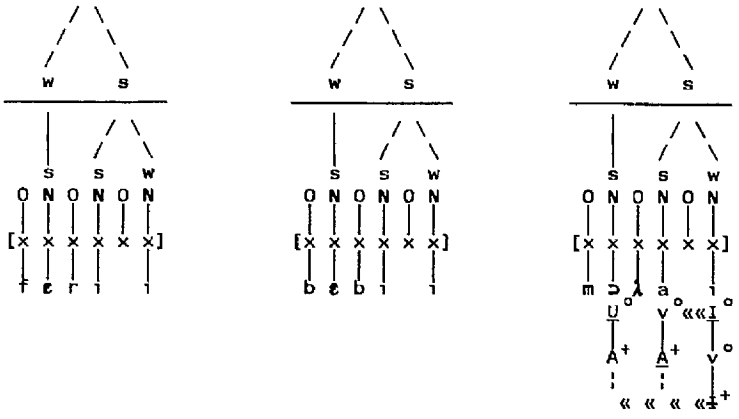
P6



[fìrírɑ̌ɯ] [bèbérɑ̌ɯ] [mɔ̀lárɑ̌ɯ]

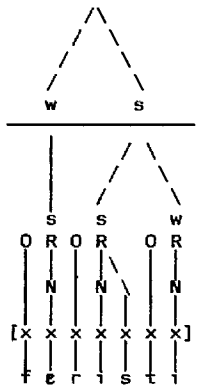
(4) The Indicative Past.

P1

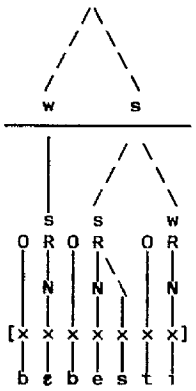


[fìríĭ] [bèbíĭ] [mòléĭ]

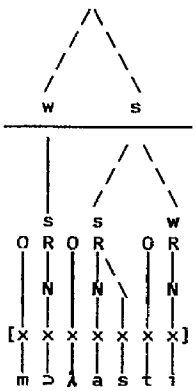
P2



[fɪrɪstɪ̌]

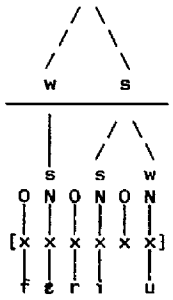


[bɛ̀bɛ́stɪ̌]

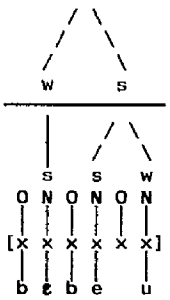


[mɔ̀ʌ́stɪ̌]

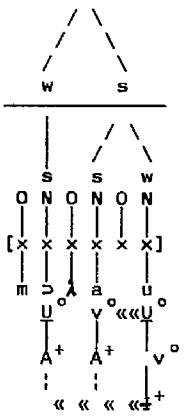
P3



[fɪrɪ́ŭ]

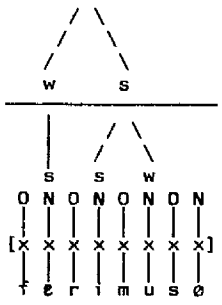


[bɛ̀bɛ́ŭ]

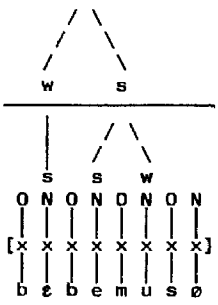


[mɔ̀ʌ́ŭ]

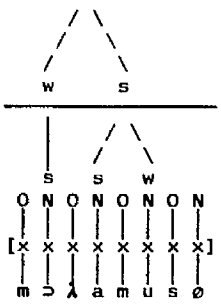
P4



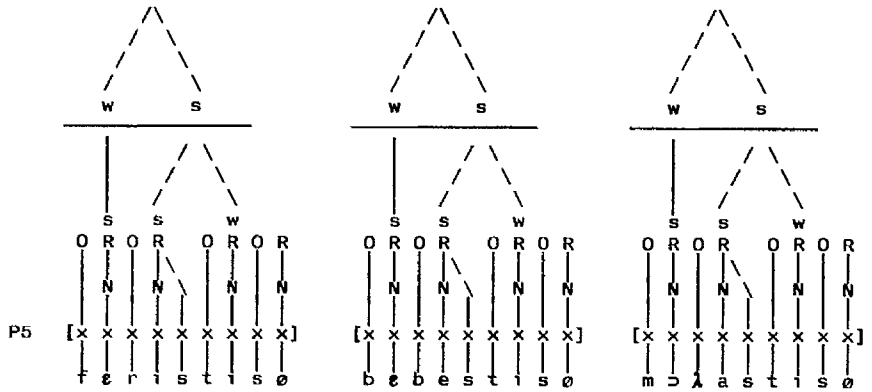
[fɪrɪ̃mŭs]

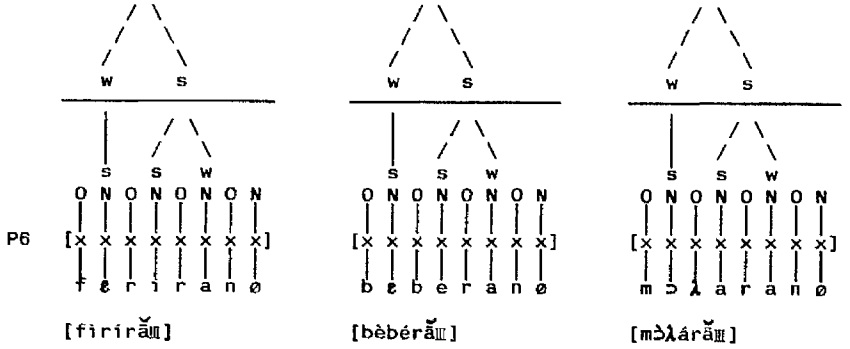


[bɛ̀bɛ̃mŭs]

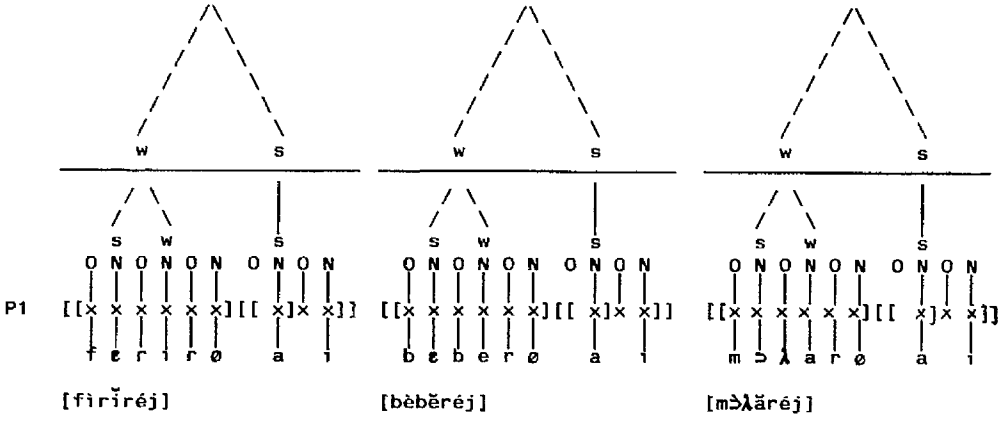


[mɔ̀ʌ́ãmŭs]

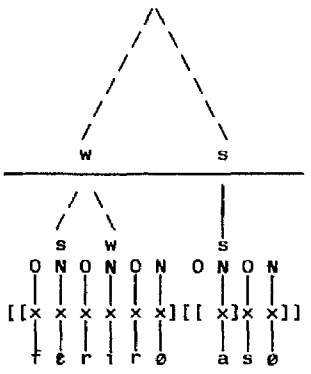

P5
[fìrístĭs]
[bèbéstĭs]
[mɔ̀lástĭs]


P6
[fìrírăɯ]
[bèbérăɯ]
[mɔ̀lárăɯ]

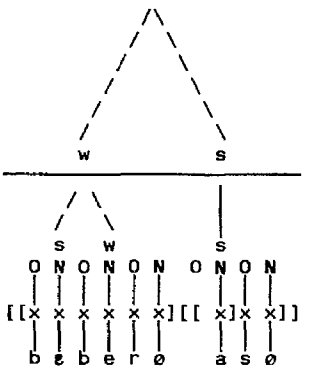
(5) The Future


P1
[fìrĭréj]
[bèbĕréj]
[mɔ̀lăréj]

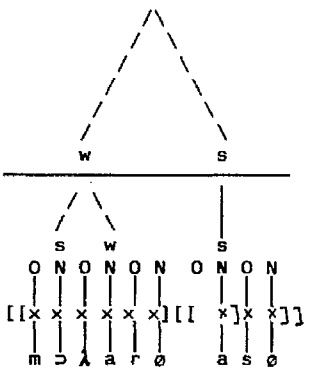
P2



[fìrĭrájs]

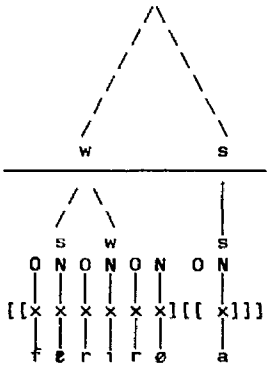


[bèbĕrájs]

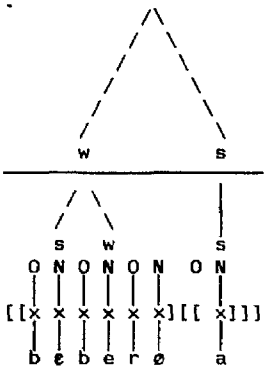


[mɔ̀lărájs]

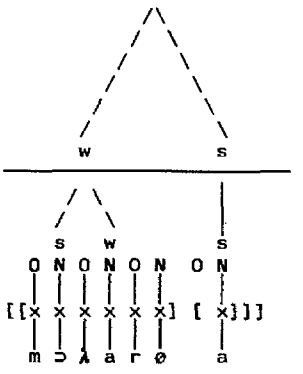
P3



[fìrĭrá]

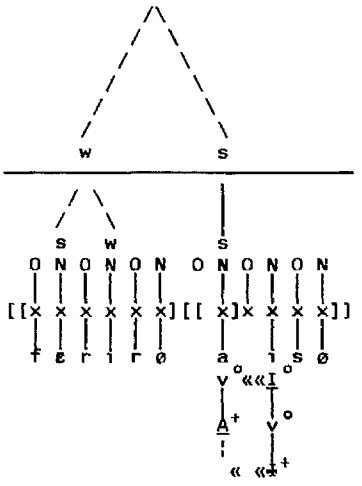


[bèbĕrá]

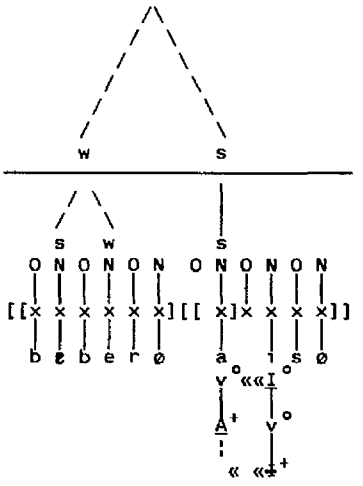


[mɔ̀lărá]

P5

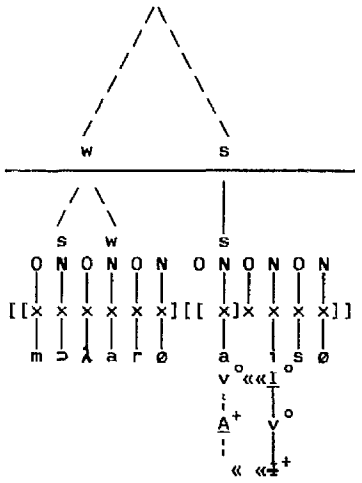


[fìrĭréĭs]



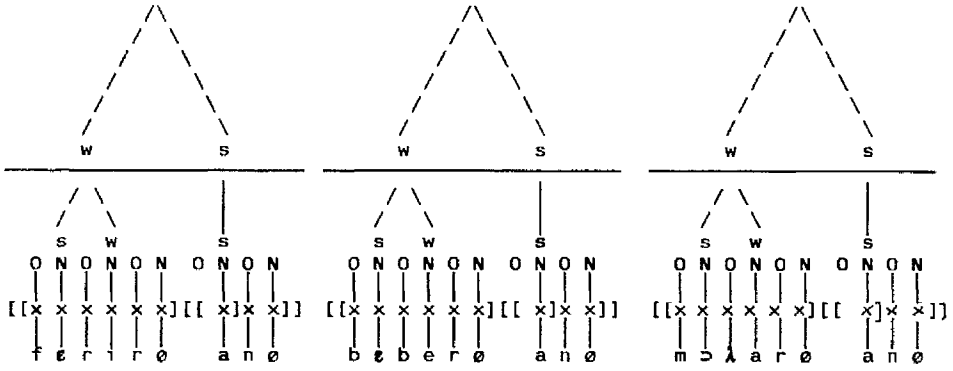
[bèbĕréĭs]

P5



[mɔ̀ɬărέĭs]

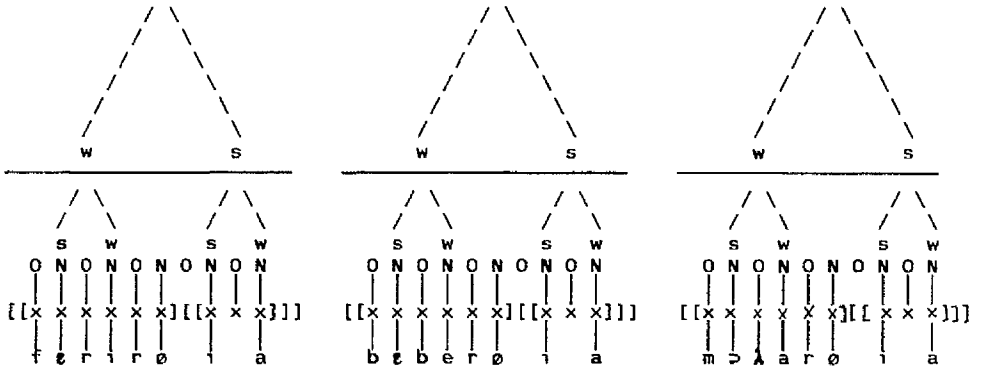
P6



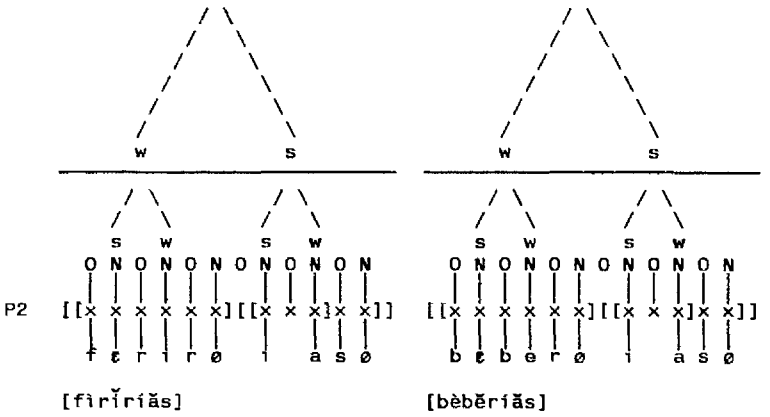
[fìrĭrầɯ] [bèbĕrầɯ] [mɔ̀ɬărầɯ]

(6) The Conditional.

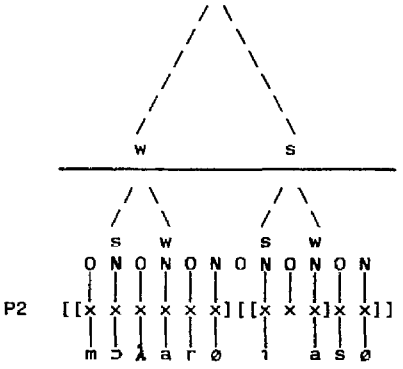
P1



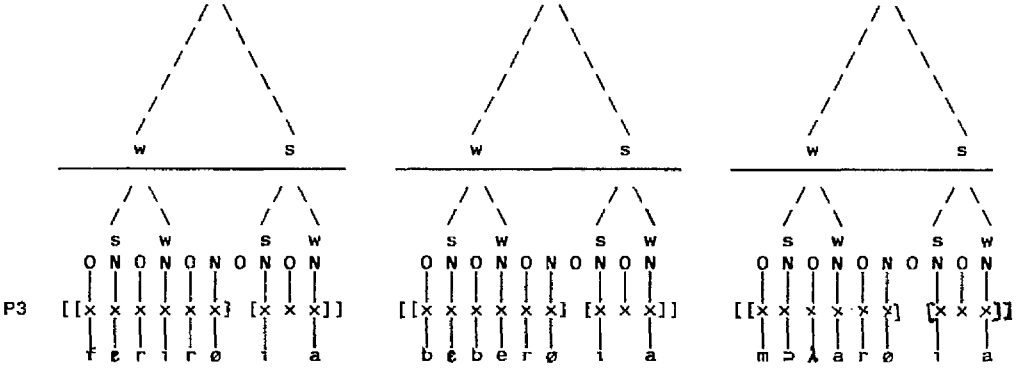
[fìrĭrìă] [bèbĕrìă] [mɔ̀ɬărìă]



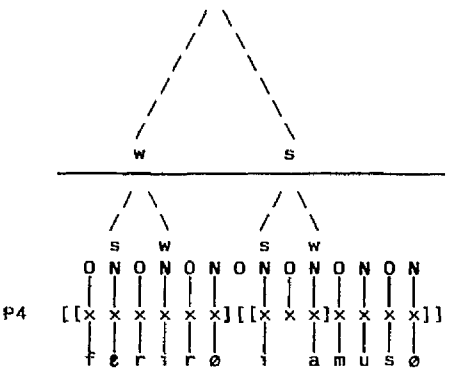
[fɪrĭrɪăs] [bèbĕrɪăs]



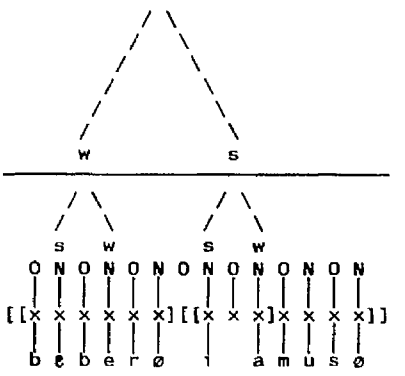
[mɔ̀ʎărɪăs]



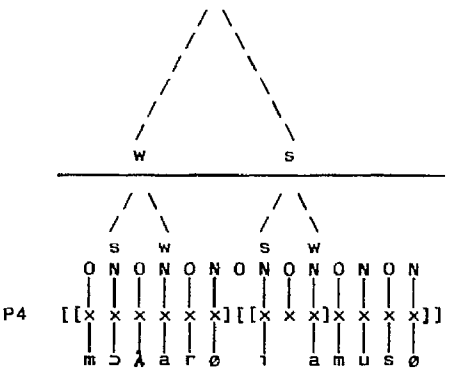
[fɪrĭrɪă] [bèbĕrɪă] [mɔ̀ʎărɪă]



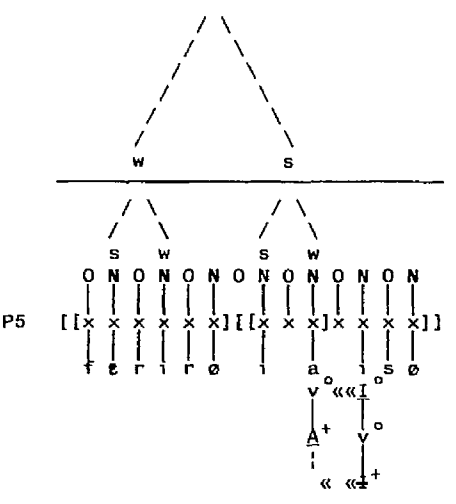
P4 [fìrĭríăm̃ŭs]



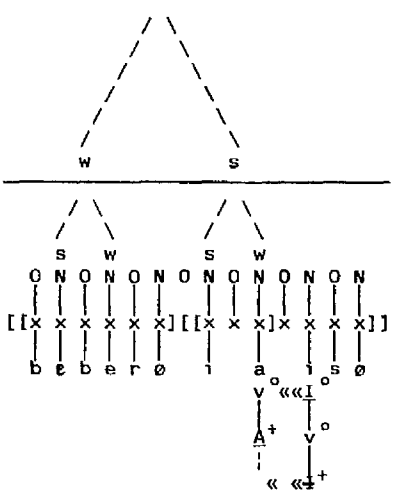
[bèbĕríăm̃ŭs]



P4 [mɔ̀ʌ̀ărĭăm̃ŭs]

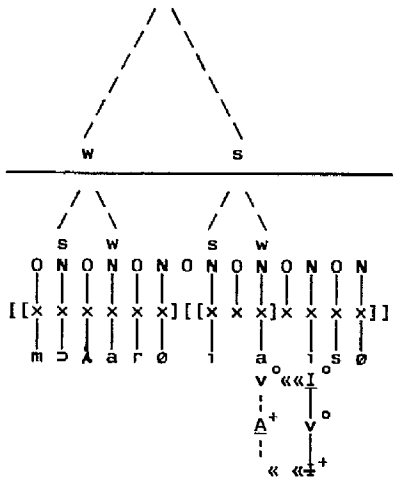


P5 [fìrĭríĕĭs]



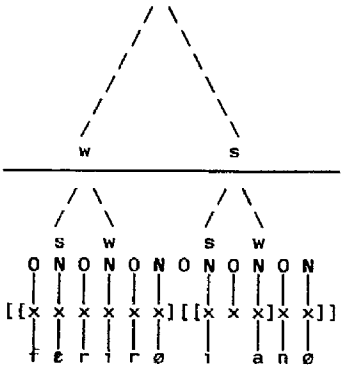
[bèbĕríĕĭs]

P5

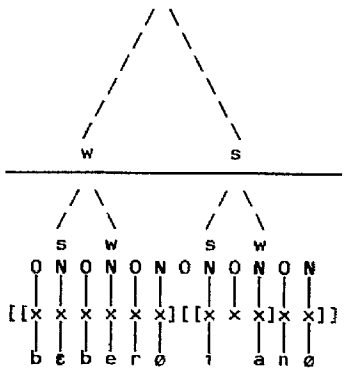


[mɔ̀ɬărɪ́ĕ̌ɪs]

P6

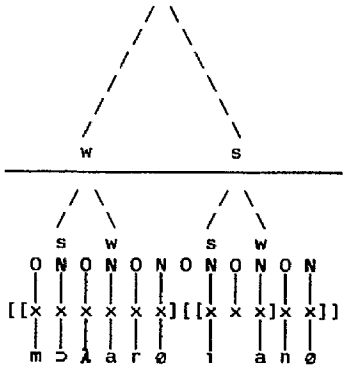


[fɪ̀rɪ̌rɪ́ă᷄ɯ]



[bèbĕrɪ́ă᷄ɯ]

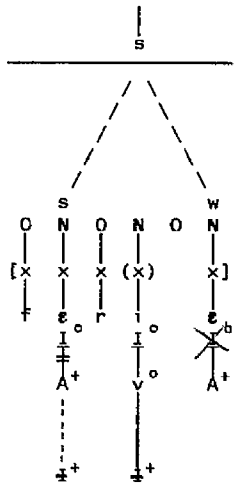
P6



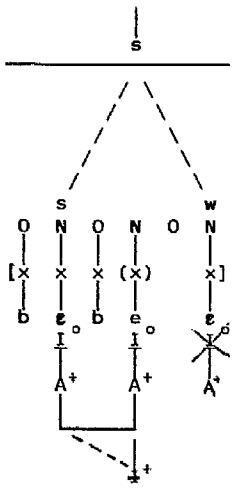
[mɔ̀ɬărɪ́ă᷄ɯ]

(7) Subjunctive Present

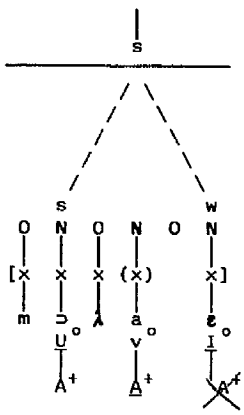
P1



[fírə̆]

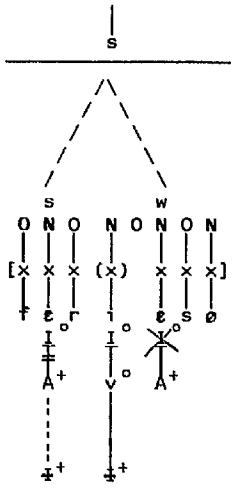


[bébə̆]

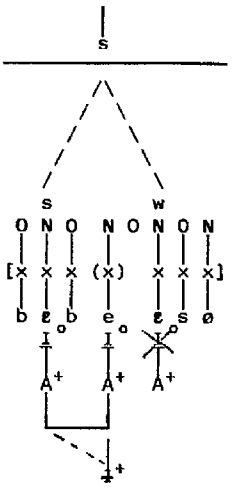


[mɔ́lɪ̆]

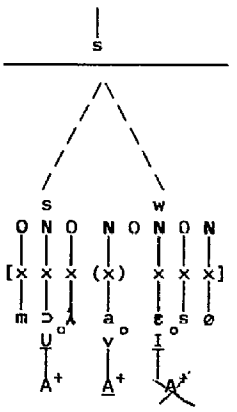
P2



[fírə̆s]

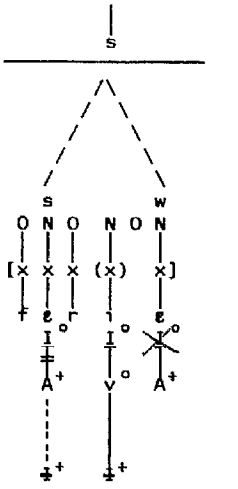


[bébə̆s]

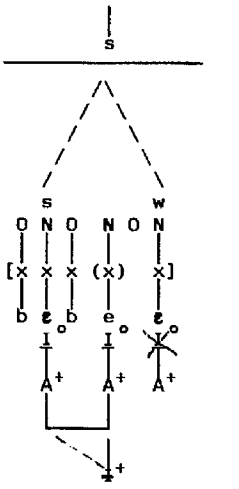


[mɔ́lɪ̆s]

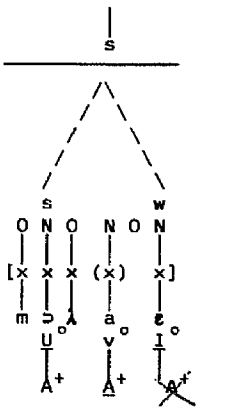
P3



[fírə̆]

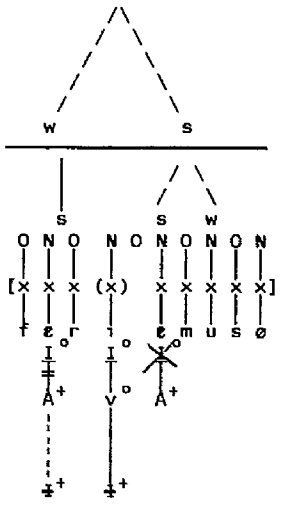


[bébə̆]

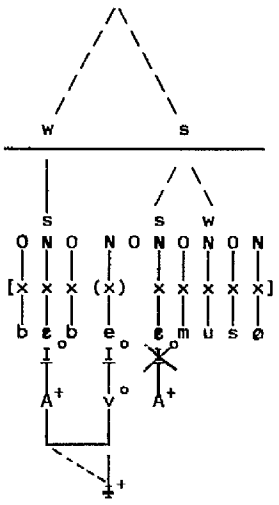


[mɔ́lɪ̆]

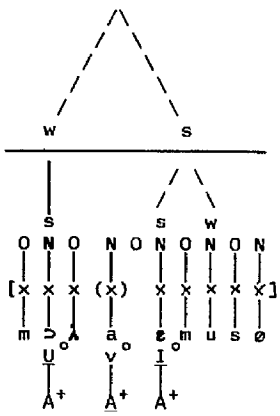
P4



[fìrãmŭs]

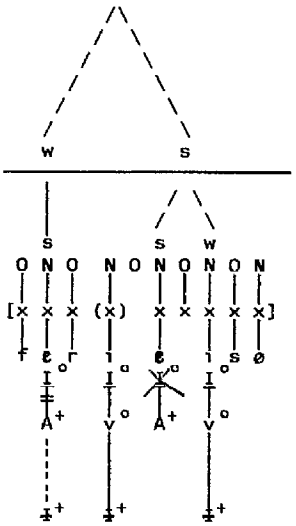


[bèbãmŭs]

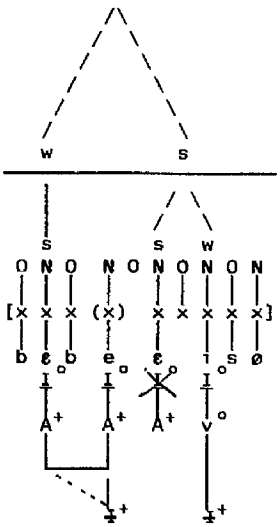


[mɔ̀ʌ̃ẽmŭs]

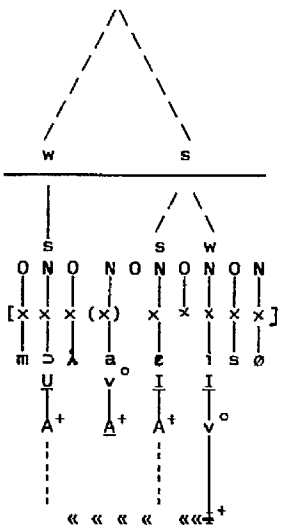
P5



[fìráĭs]

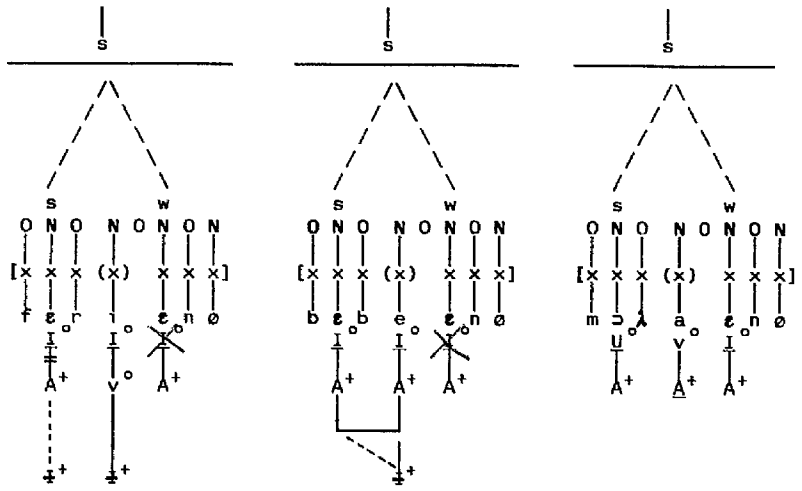


[bèbáĭs]



[mòʌéĭs]

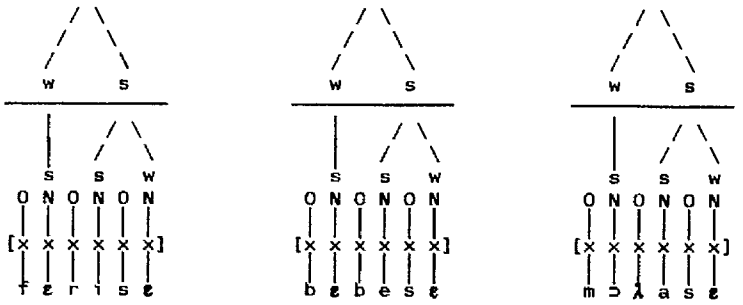
P6



[firãw] [bébãw] [mɔ́lɛ̃j]

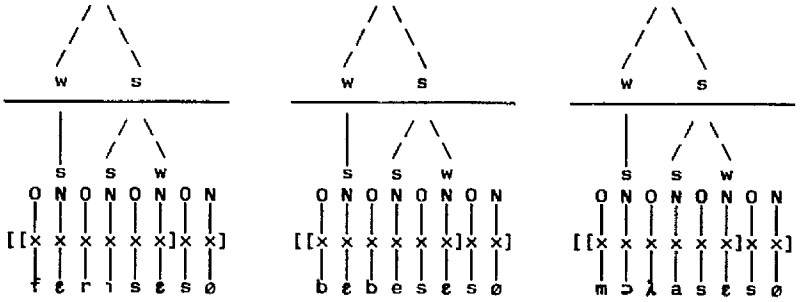
(8) The Subjunctive Imperfect.

P1



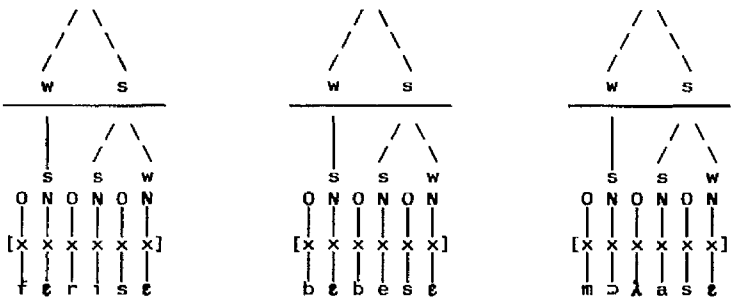
[firísĭ] [bèbésĭ] [mɔ̀lásĭ]

P2



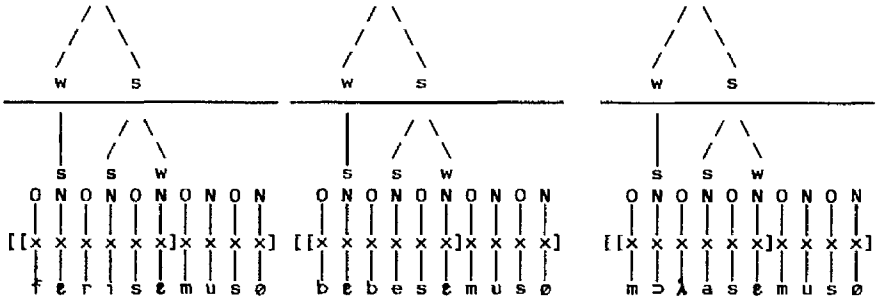
[firísĭs] [bèbésĭs] [mɔ̀lásĭs]

P3



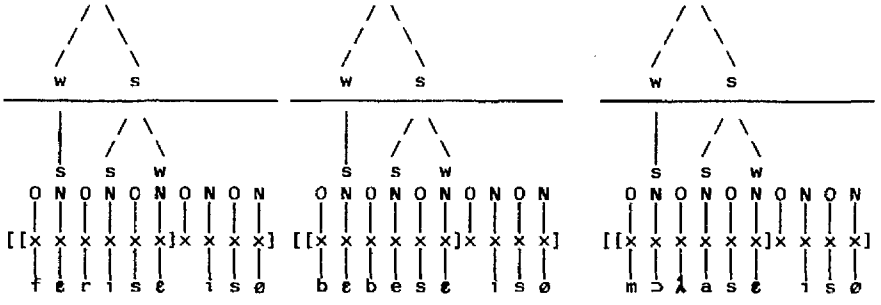
[fìrísǐ] [bèbésǐ] [mɔ̀ʎásǐ]

P4



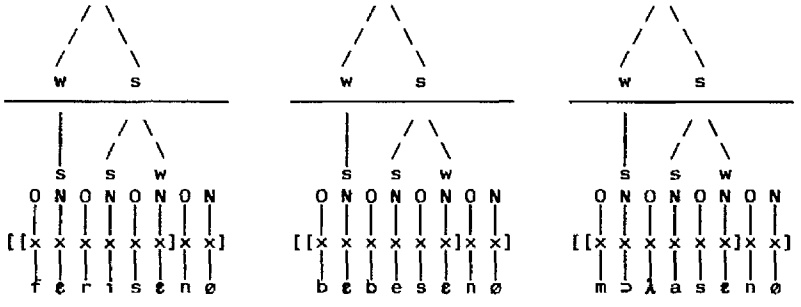
[fìrìsǐmŭs] [bèbésǐmŭs] [mɔ̀ʎásǐmŭs]

P5



[fìrísěǐs] [bèbésěǐs] [mɔ̀ʎásěǐs]

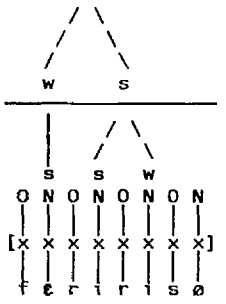
P6



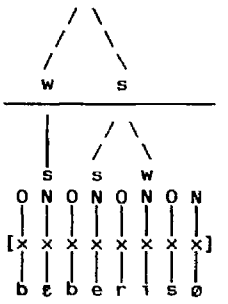
[fìrísɛ̌j] [bèbésɛ̌j] [mɔ̀ʎásɛ̌j]

(9) The Subjunctive Future.

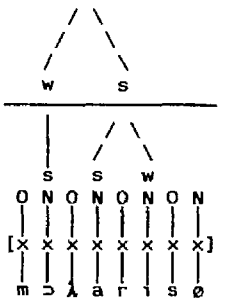
P2



[fìrírĭs]

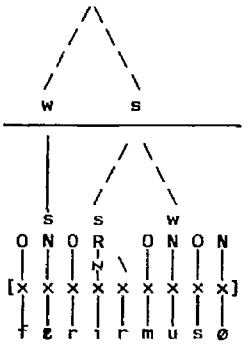


[bèbérĭs]

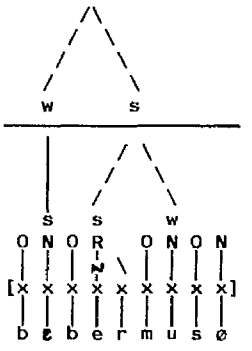


[mɔ̀ɲárĭs]

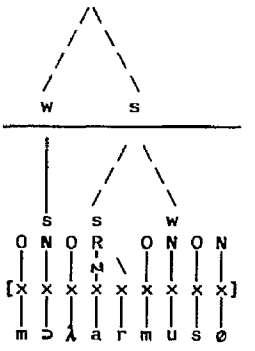
P4



[fìrímmũs]

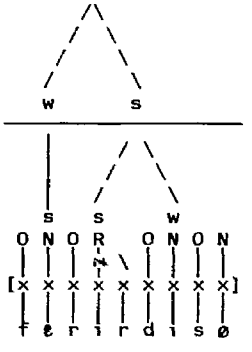


[bèbémmũs]

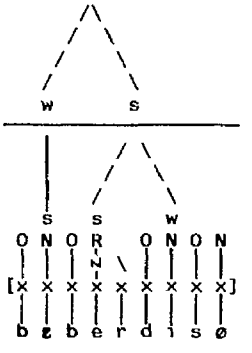


[mɔ̀ɲámmũs]

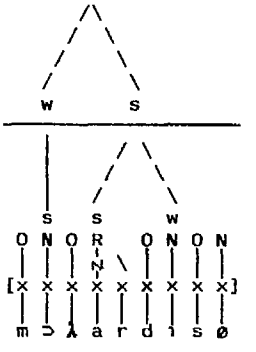
P5



[fìrídzĭs]

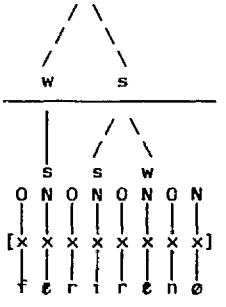


[bèbédzĭs]

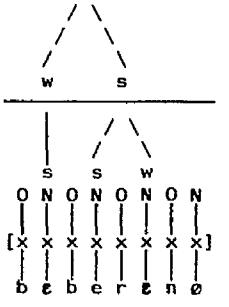


[mɔ̀ɲádzĭs]

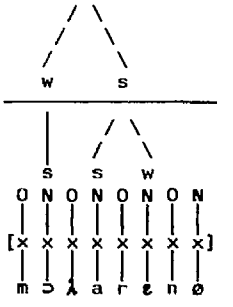
P6



[fìrírɛ̃j]



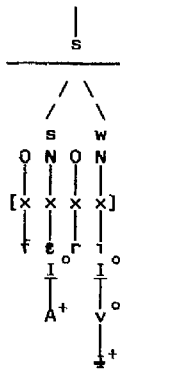
[bèbérɛ̃j]



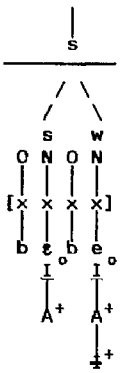
[mɔ̀ɲárɛ̃j]

(10) The Affirmative Imperative.

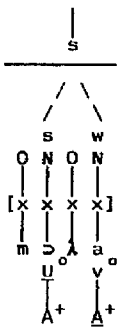
P2a



[férĭ]

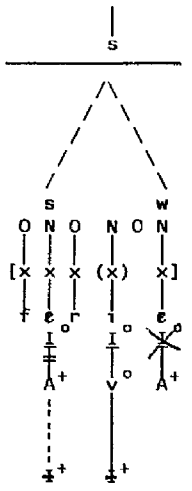


[bébĭ]

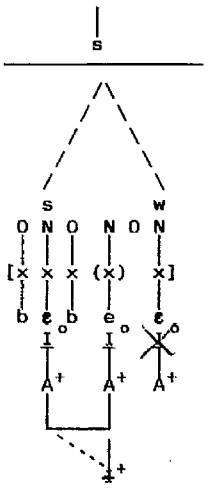


[mɔ́lă]

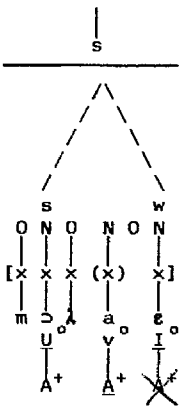
P2b



[fíră]

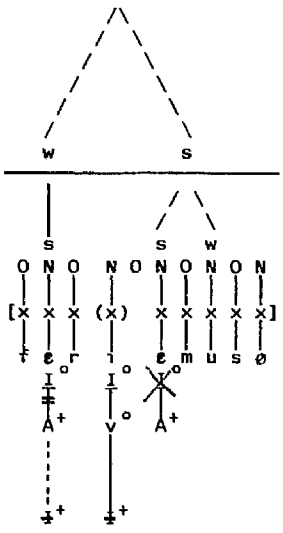


[bébă]

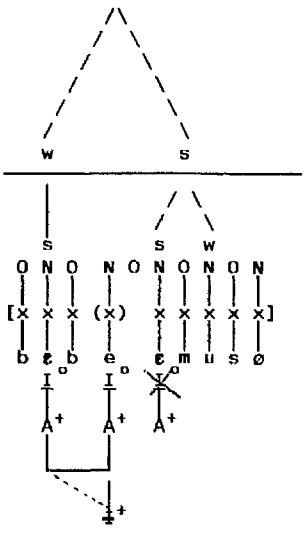


[mɔ́lĭ]

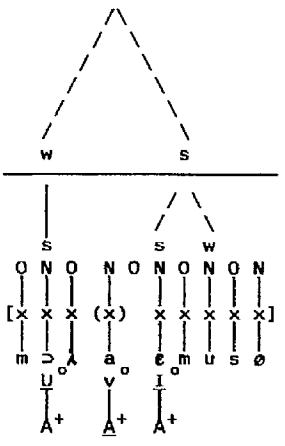
P4



[fìrə̃mŭs]

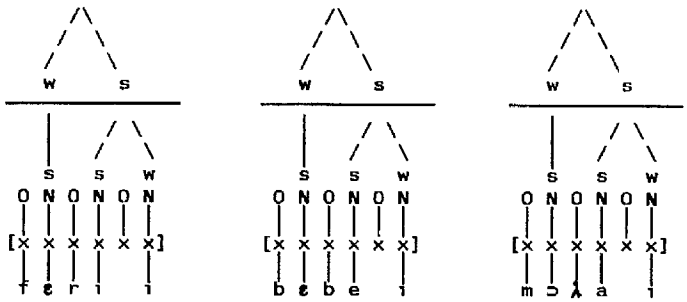


[bèbə̃mŭs]



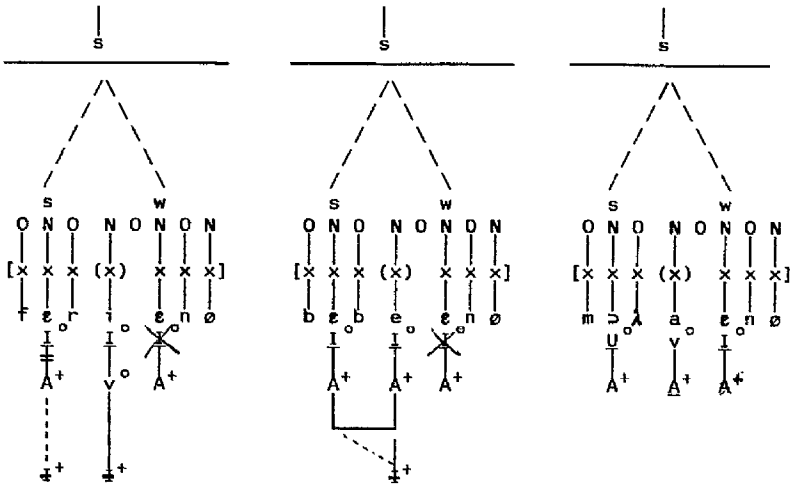
[mɔ̀lɛ̃mŭs]

P5



[fìrìǐ] [bèbéǐ] [mòɬáǐ]

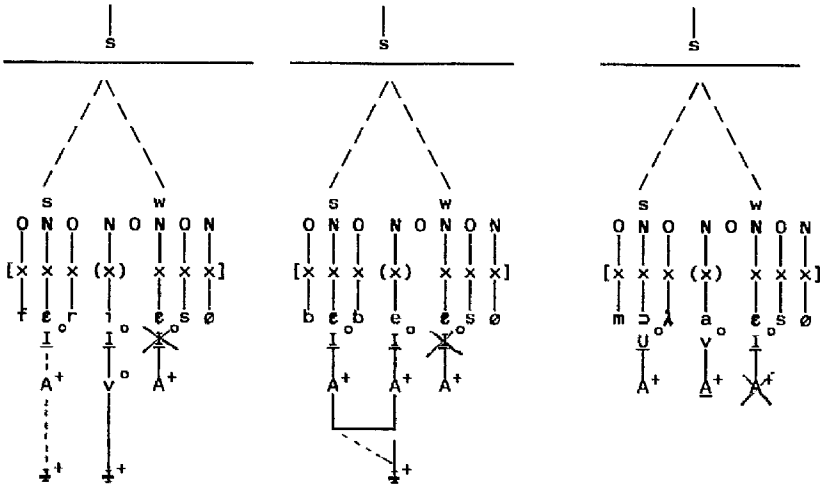
P6



[fírə̆w] [bébə̆w] [mɔ́ɬɛ̆j]

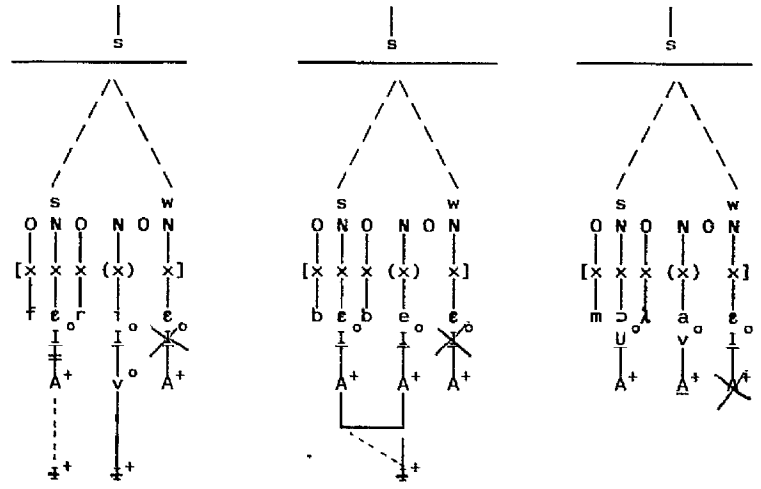
(11) The Negative Imperative

P2a



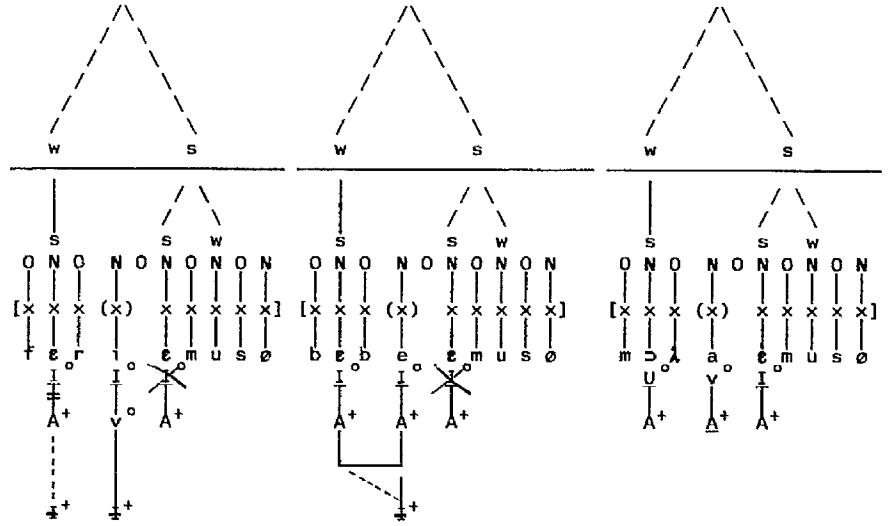
[fírə̆s] [bébə̆s] [mɔ́ɬɪ̆s]

P2b



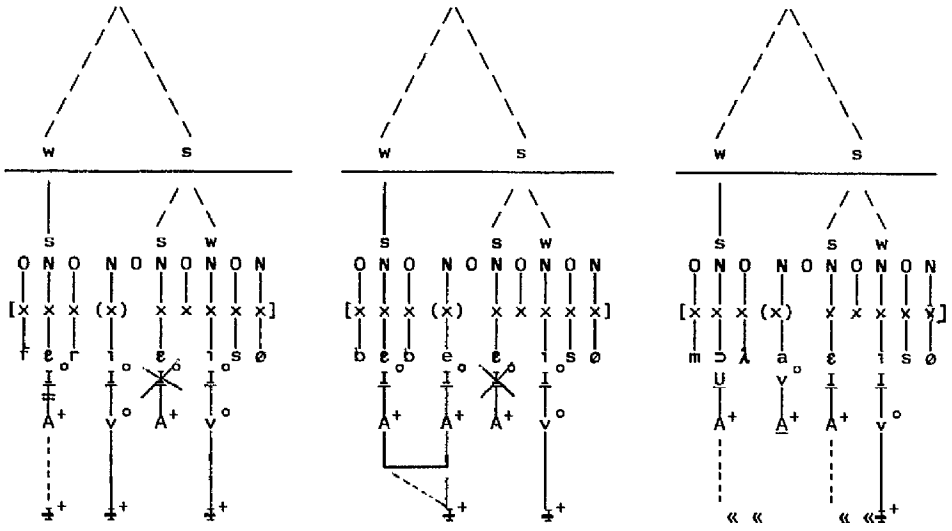
[firă] [bébă] [mɔ́λĭ]

P4



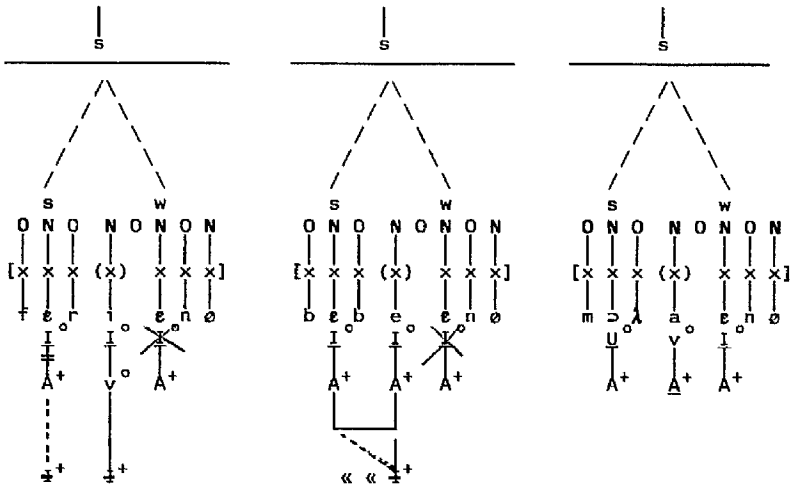
[fì rã́mŭs] [bèbã́mŭs] [mɔ̀λɛ́mũs]

P5



[fìráǐs] [bèbáǐs] [mòʎéǐs]

P6



[fíráw̌] [bébáw̌] [mɔ́ʎéǰ]

References

BUESCU, M.L.C. (1984) *Historiografia da Língua Portuguesa*, Livraria Sá da Costa Editora, Lisbon, Portugal.

CÂMARA, (1975) (Cf. MATTOSO CÂMARA 1975)

CHARETTE, M. (1990) "Licence to Govern", in: *Phonology*, vol. 7.2: 233-253.

COUTINHO, I.L. (1977) *Gramática Histórica*, 7th ed., Ao Livro Técnico, Rio de Janeiro, Brasil.

DA SILVA, T.C. (1988) *The Diminutive in Brazilian Portuguese*, London Phonology Seminar, SOAS, University of London.

DA SILVA, T.C. (1992) *Nuclear Phenomena in Brazilian Portuguese*, PhD Thesis, SOAS, University of London.

FUDGE, E. (1973), *Phonology: Selected Readings* (ed.), Penguin.

GARDE (1973) "Principles of the Synchronic Description of Stress", in: *Phonology*, E. Fudge (ed.), Penguin, p.309-319.

HALLE, M. & J-R. VERGNAUD (1987) *An Essay on Stress*, MIT Press, U.S.A.

HARRIS, J.W. (1983) *Syllable Structure and Stress in Spanish: A Non-Linear Analysis*, U.S.A., MIT Press.

HOGG, R. & C.B. McCULLY (1987) *Metrical Phonology: A Coursebook*, C.U.P.

KAYE, J.D. (1989) *Phonology: A Cognitive View*, Lawrence Erlbaum Associates Publishers, London.

KAYE, J.D. (1990a) "Government in Phonology: The case of Moroccan Arabic", *The Linguistic Review*, 6:131-160, "1986/1987".

KAYE, J.D. (1990b) "'Coda' Licensing", in: *Phonology*, vol. 7.2:301-330.

KAYE, J.D. (1992) "On the interactions of theories of Lexical Phonology and theories of phonological phenomena", in: *Phonologica 1988*, W.U. Dressler, H.C. Luschützky, O.E. Pfeiffer, and J.R. Rennison (eds.), CUP, pp.141-155.

KAYE, J.D. (to appear) "Derivations and Interfaces", to appear in DURAND, J. & F. KATAMBA, *New Frontiers in Phonology*, Harlow, Essex, Longman.

KAYE, J.D., J. LOWENSTAMM & J-R VERGNAUD (1985) "The Internal Structure of Phonological Elements: A Theory of Charm and Government", in: *Phonology Yearbook*, vol. 2: 305-328.

KAYE, J.D., J. LOWENSTAMM & J-R VERGNAUD (1990) "Constituent Structure and Government in Phonology ", in: *Phonology* 7.2: 193-231.

KEAN, M-L (1974) "The Strict Cycle in Phonology", in: *Linguistic Inquiry*, vol. 5.2: 179-203.

LEBEN (1973) *Suprasegmental Phonology*, PhD Dissertation, MIT.

LIBERMAN, M. & A. PRINCE (1977) "On Stress and Linguistic Rhythm", in: *Linguistic Inquiry*, vol. 8: 249-336.

MACEDO, W. (1979) *Gramática de Ouro da Língua Portuguesa: Guia Prático de Fonologia, Morfologia e Sintaxe*, Tecnoprint, Rio de Janeiro, Brasil.

MAIA, E.A.M. (1981) *Hierarquias de Constituentes em Fonologia.* Anais do V Encontro Nacional de Lingüística, PUC-RJ, p. 260-289.

MAJOR, R. (1981) "Stress-Timing in Brazilian Portuguese", in: *Journal of Phonetics*, 9:343-51.

MAJOR, R. (1985) "Stress and Rhythm in Brazilian Portuguese", in: *Language*, Vol. 61.2: 259-282.

MATEUS, M.H.M. (1982) *Aspectos de Fonologia Portuguesa*, INIC, Lisbon, Portugal.

MATEUS, M.H.M. (1983) "O Acento de Palavra em Português: Uma Nova Proposta", in: *Boletim de Filologia*, vol. I: 211-229.

MATEUS, M.H.M., A.M. BRITO, I. DUARTE, & I.H. FARIA (1989), *Gramática de Ouro da Língua Portuguesa*, 2nd ed.,Caminho, Lisbon, Portugal.

MATTOSO CÂMARA Jr, J. (1975) *Estrutura da Língua Portuguesa*, 6th ed., Vozes, Rio de Janeiro, Brasil.

MATTOSO CÂMARA Jr., J. (1986) *Problemas de Lingüística Descritiva*, 12th ed., Vozes, Rio de Janeiro, Brazil.

NESPOR, M. & I. VOGEL (1988) On Lapses and Clashes, in: *Phonology*, vol. 6, CUP, United Kingdom, pp. 69-116.

NESPOR, M. & I. VOGEL (1989) *Prosodic Phonology*, Foris Publications, Dordrecht, The Netherlands.

PARKINSON, S. (1988) "Portuguese", in: HARRIS, M. & N. VINCENT, *The Romance Languages*, Routledge, London, United Kingdom, p. 131-169.

PRINCE (1983) "Relating to the Grid", in: *Linguistic Inquiry*, Vol. 14.1: 19-100.

RYAN, M.A. (1988) *Conjugação de Verbos em Português*, 4th ed., Ática, São Paulo, Brasil.

SELKIRK, E.O. (1984) *Phonology and Syntax: The Relation Between Sound and Structure*, MIT Press, U.S.A.

VERGNAUD, J-R. (1982) *On the Foundations of Phonology*, paper presented at the 1982 GLOW Colloquium, Paris.

YOSHIDA, S. (1991) *Some Aspects of Governing Relations in Japanese Phonology*, PhD Dissertation, University of London.